The role of water in the kinetics of hydrophobic molecular recognition investigated by stochastic modeling and molecular simulations

Dissertation

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM

(Dr. rer. nat.)

im Fach Physik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
von

Herrn M.Sc. Richard Gregor Weiß

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Elmar Kulke

1. Gutachter: Prof. Dr. Joachim Dzubiella

2. Gutachter: Prof. Dr. Roland R. Netz

3. Gutachter: Prof. Dr. Benjamin Lindner

Tag der mündliche Prüfung: 1. Februar 2018

Richard Gregor Weiß

The role of water in the kinetics of hydrophobic molecular recognition investigated by stochastic modeling and molecular simulations

Dissertation zur Erlangung des akademischen Grades

DOKTOR RERUM NATURALIUM (Dr. rer. nat.)

Betreuer: Prof. Dr. Joachim Dzubiella

List of publications

This work comprises the content of the following original papers:

[I] R. G. Weiß, P. Setny and J. Dzubiella "Solvent fluctuations induce non-Markovian kinetics in hydrophobic pocket-ligand binding". *J. Phys. Chem. B* 120.33 (2016), pp. 8127–8136

[II] R. G. Weiß, P. Setny and J. Dzubiella "Principles for tuning hydrophobic ligand-receptor binding kinetics". *J. Chem. Theory Comput.* 13.6 (2017), pp. 3012–3019

[III] R. G. Weiß, R. Chudoba, P. Setny and J. Dzubiella "Affinity, kinetics, and pathways of anisotropic ligands binding to hydrophobic model pockets". *In preparation*

Other publications of the author that are not included in this thesis:

[IV] B. Medetov, R. G. Weiß, Z. Z. Zhanabaev and M. A. Zaks "Numerically induced bursting in a set of coupled neuronal oscillators". *Commun. Nonlinear Sci. Numer. Simul.* 20.3 (2015), pp. 1090–1098

[V] R. G. Weiß, M. Heyden and J. Dzubiella "Curvature dependence of hydrophobic hydration dynamics". *Phys. Rev. Lett.* 114.18 (2015), 187802

[VI] A. Phan, D. R. Cole, R. G. Weiß, J. Dzubiella and A. Striolo "Confined water determines transport properties of guest molecules in narrow pores". *ACS Nano* 10.8, pp. 7646–7656 (2016)

[VII] P. Bodemeyer, R. G. Weiß, S. Angioletti-Uberti and J. Dzubiella "The role of charge regulation in ligand binding kinetics". *In preparation*

Abstract

The association of small molecules (ligands) to hydrophobic binding pockets plays an integral role in biochemical molecular recognition and function, as well as in various self-assembly processes in the physical chemistry of aqueous solutions. The binding process is typically governed by the “key-lock” principle, in which the shape-complementary pocket and ligand associate due to intrinsic and water-mediated interactions. While the investigation of water contributions to the binding free energy (affinity) in equilibrium has attracted a great deal of attention in the last decade, little is known about the role of water in determining the rates of binding and kinetic mechanisms. For instance, what are the nanoscale water effects on ligand diffusion close to the hydrophobic docking site, and how can they be steered by the chemical composition of the pocket?

Recent studies used molecular simulations of a simple prototypical pocket-ligand model to show that hydration fluctuations within the binding pocket can couple to the ligand dynamics and influence its binding rates. Since the hydration fluctuations, in turn, can be modified by the pocket’s geometry and hydrophobicity, the possibility exists to create well-controlled solvent fluctuations to steer the ligand’s binding rates. In this work, we pick up this appealing notion employing a theoretical multi-scale approach of a generic key-lock system in aqueous solution utilizing stochastic dynamical systems and molecular dynamics simulations. In particular, we rationalize the kinetic coupling of hydration fluctuations and ligand dynamics regarding the fluctuation-dissipation theorem. We explore the influence of the physicochemical properties of the pocket on local ligand diffusivities and binding rates and demonstrate how the orientation of a (non-spherical) ligand couples to a pocket’s hydration fluctuations. We find that minor modulation in pocket depth can drastically speed up the binding rate and that, concurrently, binding to molded binding sites is advantageous for the rotational dynamics of the ligand. Throughout we elaborate on the methods for key-lock binding kinetics adding to our fundamental understanding of (bio)molecular processes far from equilibrium. The results and discussion of this work shall, therefore, imply generic design principles for tailored solutions of functional host-guest systems as well as optimized drugs in biomedical applications.

Zusammenfassung

Die Assoziation kleiner Moleküle (Liganden) in hydrophobe Bindungstaschen spielt eine fundamentale Rolle in der Biomolekularerkennung und den Selbstassemblierungsprozessen der physikalischen Chemie wässriger Lösungen. Der Bindungsprozess unterliegt typischerweise dem sog. „Schlüssel-Schloss“-Prinzip, in dem die komplementären Geometrien der Bindungspartner durch intrinsische und wasserinduzierte Wechselwirkungen zusammengeführt werden. Während der Einfluss des Wassers auf die freie Energie der Bindung (die Bindungsaffinität) im thermischen Gleichgewicht in den letzten Jahren auf immer stärkere Aufmerksamkeit stößt, ist die Rolle des Wassers in der Kinetik und der Bestimmung der Bindungsraten noch weitestgehend unverstanden. Welche nanoskaligen Effekte des Wassers beeinflussen die Dynamik des Liganden in der Nähe der Bindungstasche, und wie lassen sie sich durch die chemischen Eigenschaften der Tasche steuern?

Neuste Forschungen haben mithilfe von molekularen Computersimulationen eines einfachen Modells gezeigt, dass Hydrationsfluktuationen in der hydrophoben Bindungstasche enorm an die Dynamik des Liganden koppeln und damit seine Bindungsrate beeinflussen. Da die Wasserfluktuationen wiederum durch die Geometrie und Hydrophobizität der Bindungstasche beeinflusst werden, entsteht die Möglichkeit, kontrollierte Fluktuation zu kreieren, um die Bindungsraten des Liganden zu steuern. In dieser Arbeit wird diese interessante Perspektive mithilfe eines theoretischen Multiskalenansatzes für prototypische Schlüssel-Schloss-Systeme aufgegriffen. Dafür werden Computersimulationen sowie ein stochastisches Diffusionsmodell verwendet. Im Speziellen berechnen wir die kinetische Kopplung mithilfe des Fluktuation-Dissipation-Theorems. Wir untersuchen systematisch den Einfluss der physikochemischen Eigenschaften der Bindungstasche auf die Diffusivität und die Bindungsraten des Liganden, und wie die Orientierung eines anisotropen Liganden an die Hydrationsfluktuationen der Tasche koppelt. Damit stellen wir fest, dass kleine Änderungen der Taschentiefe eine extreme Beschleunigung der Bindungsraten bewirken kann und, dass gleichzeitig die Bindung in konkave Taschen vorteilhaft für die Reorientierungsdynamik des Liganden ist. Im Verlauf der Arbeit halten wir die Diskussion über die kinetische Methodik für Schlüssel-Schloss-Bindung aufrecht, was zu unserem fundamentalen Verständnis der (bio)molekularen Prozesse fernab vom Gleichgewicht beiträgt. Die Resultate dieses Projekts sollen somit helfen, maßgeschneiderte Lösungen für funktionale „Host-Guest“-Systeme sowie pharmazeutische Moleküle in biomedizinischen Anwendungen zu entwickeln.

Contents

1 Introduction – the solvent of terrestrial life

> *You cannot cross the sea merely by standing and staring at the water.*
>
> **— Rabindranath Tagore**
> (Polymath and Nobel laureate)

Water's eminent role on earth extends from the large scale hydrological cycle, which provides the continents with precipitation, down to the microscopic world of single molecules, in which water spans a unique network upon the constituents of living matter [1, 2]. Inferring from water's role on earth that it is the universal solvent for life would be rather speculative. Nevertheless, water strongly influences the terrestrial life – the only life we know. Scientists hypothesize that life evolved in the oceans during the early stages of our planet. Back then, water acted as an active solvent, which drove the necessary accumulation processes for the creation of life. For example, the self-assembly of lipids to micelles imposed the primary archetypes for biological cell membranes, which is a phenomenon that is very specific for lipid solvation in water. Another example is the molecular recognition and association of a substrate to a particular molded section on the surface of a folded protein where water potentially determines major contributions to the binding thermodynamics, the pathways, and the selectivity. Both examples, micelle formation and molecular recognition, are fundamental processes of molecular cell biology for which water exhibits various structural and dynamical roles [1, 3]. In this thesis, we focus on the latter, especially the binding of a non-polar ligand to a hydrophobic molded binding site. We explore the active role of water for binding kinetics in a regime where hydrodynamic continuum models are overly coarse. For that reason, let us dive down to the molecular scale of this famous solvent – water.

The single water molecule is composed of an oxygen atom that covalently binds two hydrogen atoms, as illustrated in Fig. 1.1 (a). The angle between the two hydrogen atoms is roughly 104° [4]. The electron density maximizes at the oxygen atom leaving it charged with $-2\delta \approx -0.8e$, because it is more electronegative than the hydrogen atoms. Since the whole water molecule is electroneutral, the two hydrogen atoms are also partially charged with $\delta \approx +0.4\ e$ each. The partial charges give rise to one of the strongest forces in nature, the electrostatic interaction. This

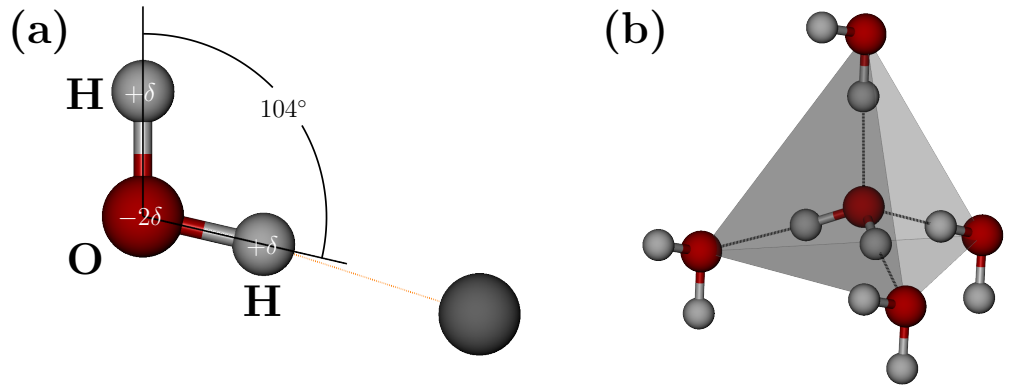

Fig. 1.1: Illustration of (a) the water molecule, which is bound to an arbitrary H-bond acceptor, and (b) the tetrahedral order of the H-bonds in ice.

force acts as a dominant intermolecular force and enables the formation of the so-called *hydrogen bonds* (H-bond): a water-hydrogen atom of one molecule can electrostatically bind to a water-oxygen atom of a neighboring molecule. In simpler terms, the former molecule is *donating* a hydrogen to the latter molecule, which is called the *acceptor* of the H-bond. (In general, the donor and acceptor need not be water molecules.) Each water molecule can donate two and accept two H-bonds, which gives a maximum of four intermolecular H-bonds on each water molecule. A tetrahedrally coordinated H-bond network spans through frozen and liquid water, as it is illustrated in Fig. 1.1 (b). In the liquid state, this molecular web fluctuates such that the single H-bond has an average lifetime on the order of a picosecond before it is intermittently ruptured and binds to another neighboring donor/acceptor [5].

In general, distortions of the H-bond network usually impose energetic penalties and thus are thermodynamically restricted. On one hand, the binding energy of a single H-bond is around 20 kJ mol^{-1}, which upon H-bond breakage, imposes an enthalpic penalty of roughly 8 k_BT at ambient conditions [6]. On the other hand, deformations by stretching and squeezing the network impose entropic penalties. The inherent minimization of all energetic penalties creates the specific properties and anomalies of water. A famous example is that water exhibits its maximal density in the liquid state at 4 °C, while the density of other substances monotonously increases with decreasing temperature. A more important example in this work is that the solvation of non-polar entities in water imposes severe penalties because a non-polar surface lacks the chemical units that could serve as H-bond acceptors or donors. Hence, water at non-polar surfaces must compromise between broken H-bonds (enthalpic costs) and network order distortions (entropic costs). Non-polar solutes are usually classified as *hydrophobic* (Attic Greek: *hýdro-* for “water” and *phóbos* for “fear”). The verbalism “hydrophobic effect” gathers the consequences

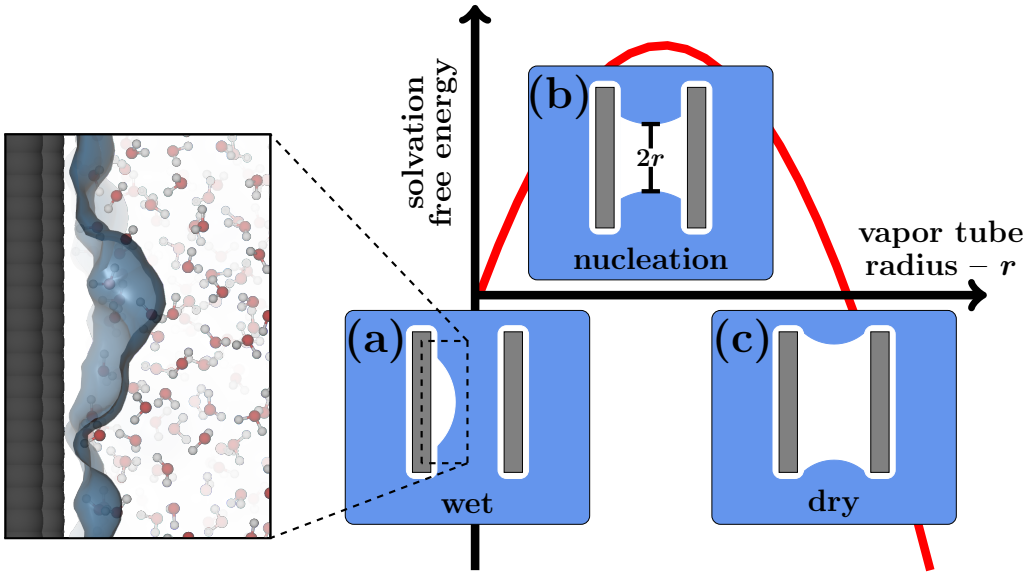

Fig. 1.2: Capillary evaporation: (a) A spontaneously formed cavity abruptly transits into (b) a cylindrical vapor tube, which grows until the confined volume is completely (c) dry. The simulation snapshot on the l.h.s shows an instantaneous water interface (blue), which exhibits a cavity.

of hydrophobicity that dictate the driving forces for macromolecular conformations and association processes like the aforementioned micelle formation and molecular recognition [7–11].

1.1 Hydrophobicity – the small and the big of it

One of the greatest advances in our understanding of hydrophobicity is the recognition that the hydration structure and thermodynamics of non-polar solutes is length scale dependent [7, 12–27]. Although the length scale dependence of solvation is not unique to water [20, 28–32], arguably it is most pronounced for water because of the self-associated or "sticky" nature due to its H-bond network. We can thus picture why the hydration thermodynamics are length scale dependent if we imagine the possible distortions of water's molecular web.

The H-bond network accommodates a small and ideally hydrophobic solute – a hard sphere – without rupturing, thus without enthalpic costs. Rather the statistics of cavity formation determine the solvation of the hard sphere which reorders, repacks, and deforms the network [33]. These distortions result in entropic costs. With increasing size, the solute dissolves while it ruptures the surrounding H-bonds because it imposes an extended steric restraint to the network [7]. In that case, the interface starts resembling a liquid-vapor like interface at which some water-hydrogens are freely dangling [7, 34]. While the interfacial water loses its self-associated nature, the molecular scale density fluctuations are enhanced and

enable spontaneous cavity formation adjacent to the surface (Fig. 1.2 (a)). Thus the surface of an extended hydrophobic solute is said to “dewet”.

If two of these extended hydrophobic objects approach each other, deceeding a critical separation, the water in between becomes metastable with respect to its vapor [14, 35–44]. A water cavitation on one of these objects can spontaneously form a cylindrical vapor tube with radius r extending from one surface to the other (Fig. 1.2 (b)), which can grow until the entire confined volume is dry (Fig. 1.2 (c)) [35–37]. The growth of the vapor tube is associated with a nucleation barrier in the solvation free energy, which is illustrated as the red line in Fig. 1.2. In the picture of macroscopic interfacial thermodynamics the barrier arises from the competition between the solid-vapor and liquid-vapor surface energies [38–43, 45]. This process is termed “capillary evaporation”, which energetically favors the dry state with decreasing separation of the non-polar objects. Moreover, at intermediate separations the (dry) volume between the two hydrophobic surfaces in Fig. 1.2 can refill with water such that the hydrophobically confined volume bimodally fluctuates from wet to dry and *vice versa*. It is important to note that the timescales of these fluctuations – the rates of crossing the nucleation barrier – are determined by the barrier height, which can be on the order of magnitude of a thermal energy k_BT.

A variety of phenomena exemplify the wide momentousness of hydrophobicity in natural, engineering and pharmaceutical sciences, such as self-assembly of amphiphilic molecules to biological membranes [7, 46], molecular recognition [9, 47–58], catalysis using cavitands [59–62], transport through nano-pores [63–65], as well as protein folding, stability, and function [66–69]. The work here especially focuses on how capillary evaporation rates influence the association kinetics of small hydrophobic ligands to hemispherically molded hydrophobic binding sites. We explore in an idealized model the implications for simple organic compounds, e.g., drugs, hormones, and toxins, which bind to biological receptors.

1.2 Molecular recognition – water’s active role

The field of molecular recognition comprises the most fundamental processes of life as well as chemically engineered processes far beyond biological matter [70]. As an example, enzymes bind a substrate to the often concavely shaped binding site and catalyze a chemical reaction of the substrate [71, 72]. Also, receptors in cell membranes are activated if their binding pockets take up a small molecule, such as a neurotransmitter [73], a hormone [74], or a pharmaceutical drug [75]. Moreover, chemical engineers copy this binding principle from nature for supramolecular chemistry, where so-called cavitands [59, 62, 76] or macrocycles [77] are designed

as molecular containers. The underlying principle is often compared to a “lock” and “key”, or a “host” and “guest”, whereas the binding agent serves as the “key” or “guest” that selectively fits into the concavely shaped “lock” or “host”. Hence, the field of molecular recognition termed the key-lock or host-guest principle [78–80] for biomolecular binding and adopted the principle for engineered molecular cavitands.

Especially the key-lock binding in water exhibits solvent-mediated effects, which in fact can be divergent [81]. On the one hand, water can provide a *plasticity* for highly adaptable binding if it serves as “brick-like” filler inside a receptor that takes up different small ligands in a voluminous binding site for highly malleable binding, which for instance is the case for oligopeptide binding protein OppA [82, 83]. On the other hand, water can provide a *specificity* if it bridges H-bonded interactions between a receptor and ligand such that complexation is energetically favored for a certain optimal binder, which is the case for the peptide-binding site of the SH2 domain of tyrosine kinase Src [84]. Naturally, a combination of water mediated plasticity and specificity can be inferred and was for example found for bacterial l-arabinose binding protein ABP for various sugar molecules [85].

1.2.1 Hydrophobicity in key-lock binding thermodynamics

A general, outstanding role is unanimously ascribed to hydrophobicity. Despite the fact that the folded configuration of a protein mostly buries its hydrophobic sections into the interior, important molded surface regions are hydrophobic and serve as vital binding motifs for hydrophobic ligands. Researchers estimate that especially the non-polar portions cover 75% of the surface area of most ligands and active sites on proteins [86]. Given that the magnitude of the free energy of binding in water scales linearly with this amount of solvent-exposed surface area suggests a fundamental role of hydrophobicity in drug discovery. Hence nature seems to exploit the motif of hydrophobicity preferentially. Investigations that account for water-mediated interactions in rational design principles for drug discovery employ empirical hydrophobicity scales, which account for the two-sided nature of hydrophobicity – surface topography and chemical composition [57, 86]. For instance, recent simulation studies indicated that tubularly *confining* protein pockets comprise stronger contributions to a ligand’s binding affinity than a shallower one, i.e., a hydrophobically less-confining binding site [56, 87]. Early experimental studies targeted to tune the binding affinity of enzyme pockets by *chemical modifications* and reported enhanced binding to increasingly hydrophobic motifs [57].

Within this framework, one argument is that ligand binding is entropically driven because the water in a hydrophobically confining receptor could cluster due to

hydrogen bonding. Hence, the water's replacement into the bulk upon ligand binding should result in an entropy dominated hydrophobic association [56, 88]. This point only holds true in some cases. In contrast, the occupying water can also be highly disordered in a hydrophobically confining pocket which gives increased entropy to these water molecules [53, 89]. If a ligand binds, this entropy is then lost, and the hydrophobic key-lock association is driven by the enthalpic contributions of H-bond reformation upon the release of water into the bulk. In sum, the influence of hydrophobicity on binding affinity is tunable due to chemical and topographical modifications of the binding partners [58, 90–95]. With that in mind, especially synthetic key-lock systems, i.e., cavitands, proved to be amenable and adaptable to hands-on, guided mutations [61, 62].

1.2.2 Hydrophobicity in key-lock binding kinetics

A word of caution is due: Our growing understanding of the physical phenomena in molecular recognition recommends that efficacy and safety of a chemical agent in open, *in vivo* systems are highly influenced not only by thermodynamics, i.e., binding affinity, to its target but also by its kinetic rates [96–99]. Accordingly, association and dissociation rates determine how quickly a drug/toxin binds, and how long it remains bound to a receptor, in sum its efficacious time frame. Thus, the binding affinity, described by the binding constant K_a, alone does not suffice as sole parameter for tailored drug development. We rather have to keep in mind that the binding constant, thus affinity, relates to the ratio of the association rate k_{on} and the dissociation rate k_{off}

$$K_a = \frac{k_{on}}{k_{off}} . \tag{1.1}$$

This simple relation, however, constitutes a crucial principle for tailored key-lock solutions: Even though the ratio of k_{on}/k_{off} might stay the same, absolute changes in the respective rates might be affected by modifications of transition barriers. Especially, the dissociation rate has been shown to positively correlate with a drug's efficacy [96–99].

Novel scientific results indicate a just as important role of water in the kinetics of hydrophobic key-lock association as it takes in binding thermodynamics [100]. These results address the questions that were recently posed by Philip Ball: "How are the dynamics of water and biomolecular solutes related, and how do these dynamics influence function? How are fluctuations on different timescales and spatial scales coupled? How are the properties of water modulated at surfaces, and how do these depend on the chemical and geometric features of the surface?" [100] In particular, Setny *et al.* [55] documented a direct *kinetic coupling* between water

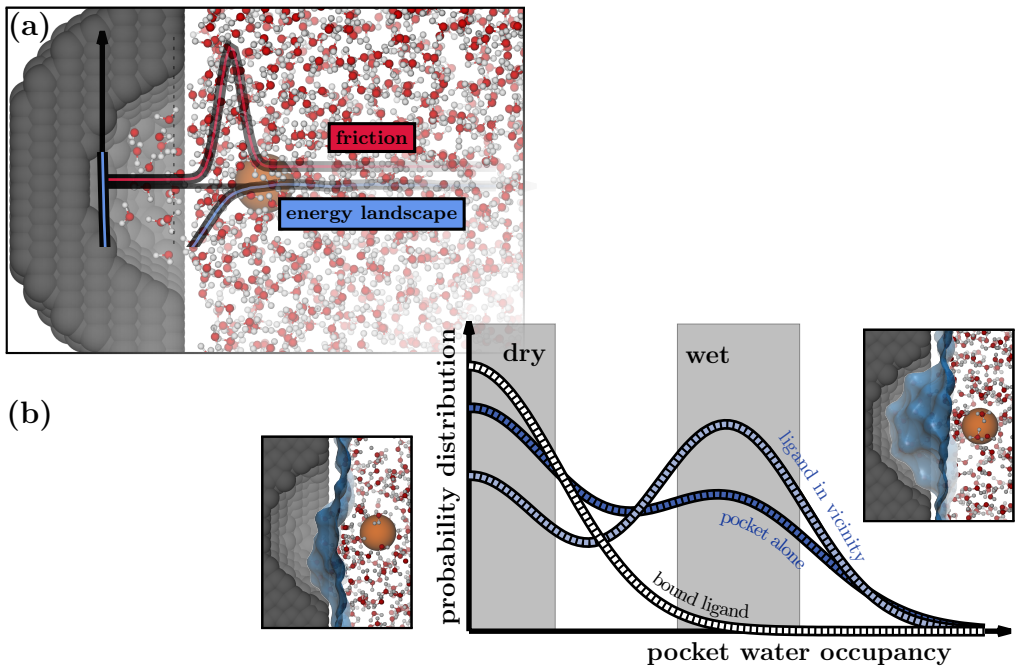

Fig. 1.3: (a) The key-lock model from Refs [48–50] comprises a hydrophobic pocket (gray) and a spherical ligand (orange) in water (red). The free energy landscape is shown as the blue line, and the local ligand friction is plotted as the red line. The friction maximizes in front of the pocket. The free energy landscape drives strong association because the hydrophobic ligand can be desolvated upon binding into the pocket. (b) If the ligand is in the vicinity of the pocket, the hydration of the pocket exhibits a bimodal distribution between dry and wet states, which are emphasized by the simulation snapshots on the left- and right-hand side of the graph. The graph itself shows different probability distributions for the pocket water occupancy of different pocket-ligand separations. These curves exemplify the bimodal hydration of the pocket.

fluctuations in hydrophobic pockets with the kinetics of an associating ligand. Parts of their findings are sketched in Fig. 1.3. They used a hydrophobic wall with a hemispherically molded pocket and a spherical ligand in explicit-water molecular dynamics (MD) simulations [48–50]. Since both, the pocket and the ligand, had no internal degrees of freedom the model served for multiple studies capturing water's pure contributions to the association (thermo)dynamics [53–55, 89]. Particularly important for this thesis, they demonstrated that bimodal hydration fluctuations due to capillary evaporation inside the pocket led to locally increased ligand friction if the ligand was in the vicinity of the binding site. They found long-time correlation effects in the ligand position and the mean force acting on it, which pointed to non-Markovian kinetic phenomena. In turn, the ligand-pocket separation also determined the hydrophobic confinement and thus the hydration fluctuations. These findings were also consistent with observations of Berne and coworkers in a study on a similar hydrophobic key-lock model setup [101], the association of two hydrophobic plates [102], or two fullerene spheres [103]. Within these studies, Berne *et al.* further

demonstrated that the enhanced friction during the association was tuned if the dispersion interaction of water with the hydrophobic objects changed. These first studies restricted the hydrophobic entities to a one-dimensional coordinate.

In subsequent studies, Tiwary and Berne [104, 105] highlighted that an orthogonal transition path in their hydrophobic key-lock model was highly unfavored. If the ligand was free to move, the coupling of ligand and pocket hydration vanished in their case. Nevertheless, these studies can only speak for the kind of model setup they used. Investigations on natural *in-vivo* systems should finally reveal where nature can employ a coupling of capillary evaporation and association kinetics. In particular contrast to Tiwary's work, a study of Mondal *et al.* [106] reported a coupling of the pocket desolvation and ligand dynamics in the special case of the kinase-inhibitor Dasatinib binding to Src-kinase. After all, we primarily motivate our study by the key-lock model from Ref. [55], and we still restrict the ligand motion to the orthogonal association pathway. Our findings still give insights on how to control and steer the possible couplings of pocket hydration fluctuations and ligand kinetics. Moreover, we elaborate on the much more general kinetic phenomena, that have implications far beyond hydrophobic key-lock models.

1.3 Outline of this work

In this dissertation, we pick up these novel findings of the kinetic coupling between a hydrophobic ligand and the hydration fluctuations in a hydrophobic binding site. In Chapter 2, we present our theoretical framework including fluctuation phenomena in statistical physics and molecular dynamics simulations. In particular, we introduce the reader to the fluctuation-dissipation theorem and explain the methods that we utilize for the MD simulation analysis. We especially elaborate on the previous findings from Refs [55, 101] whereas we depict some reproduced simulations and calculations.

In Chapter 3, we map the essential features of the hydrophobic key-lock kinetics, i.e., bimodal wet-dry fluctuations of the pocket and the ligand attraction due to its desolvation, into a minimalistic stochastic dynamical system. This part of the thesis primarily deals with a coupled (Markovian) set of stochastic differential equations in the two-dimensional coordinate space spanned by the pocket water occupancy and the ligand position. We numerically solve the dynamical equations and demonstrate how locally increased friction of the ligand, decelerated binding kinetics, and local non-Markovian (memory) effects occur in the reduced one-dimensional description alongside the ligand's distance to the binding site. Our stochastic model elucidates the origin of locally enhanced friction that can be traced back to long-time decays in

the force autocorrelation, which are induced by spatially fluctuating water-ligand interactions. Furthermore, we formulate the generalized Langevin equation of the binding process including a local memory function that reflects the dominant frequencies of the capillary evaporation in the pocket. The latter model provides further interpretation and an analytical expression of the kinetic coupling based on a fluctuation-dissipation relation.

In Chapter 4, we rationalize how the physicochemical properties of the non-polar, concave binding site tune key-lock binding kinetics in explicit-water MD simulations. We adopt the setup from Setny [48–50] to systematically modify the receptor's physicochemical properties regarding geometry and dispersion attraction, which alter the water occupancy and fluctuations in the pocket. We demonstrate that even minor pocket modifications can lead to a significant acceleration of the water-mediated association. For example, the binding switches from comparably slow to fast if the binding pocket becomes only slightly deeper. We find that the degree of hydrophobicity, quantified by pocket hydration fluctuations, clearly correlates with the binding times and, for instance, links the sudden acceleration to an abrupt increase in hydrophobicity. For a deeper analysis based on first passage time theory, we quantify the intimate coupling between the pocket hydration fluctuations and the ligand's local friction. The coupling exhibits substantial non-Markovian effects, which slow down the binding kinetics in all cases.

In Chapter 5, we explore how the hydration of a binding site affects the orientation pathway of non-spherical ligands in explicit-water simulations. We find that binding to a concavely molded binding site is advantageous for ligand orientation in a two-folded manner compared to the association with a planar site. The re-orientation process upon binding to a dewetted pocket initiates farther outside of the binding site and includes an energetic benefit. We also rescale the coordinate and the free energy by the enhanced ligand friction. This rescaling illuminates how the friction peak before binding can be interpreted as a kinetic barrier. Finally, we also discuss the unbinding which reveals that minor modifications of the ligand dominantly modulate the residence time of a substance but insignificantly influence the time until association.

In Chapter 6, we conclude the thesis. We summarize all findings to connect their implications. We discuss how our results complement the previous investigations from Setny *et al.* [52, 55] as well as from Berne and coworkers [101–105]. Further, we focus on the questions of Philip Ball mentioned above and try to give some answers, even though Ball's queries reach far beyond the scope of a single thesis. Nevertheless, we place our work into the bigger picture and interpret our results

regarding design principles for drug discovery. We depict how future work requires more insight into the modeling of capillary fluctuations at various solute surfaces and the theory of (non-Markovian) kinetics with spatially dependent long-time correlations. Finally, we propose an immediate, subsequent search for realistic and relevant receptor ligand complexes which exhibit the effects of our model system.

2 Theoretical framework – molecular statistical mechanics

Theoretical, experimental, and simulation studies explored the multifaceted topic of hydrophobicity from which we learned that a decent comprehension requires a molecular scale treatment of water [14, 107]. Also in this thesis, we mainly rely on the statistical and molecular mechanics of water while we utilize the theory of fluctuation phenomena and explicit-water MD simulations. In the first part of this chapter, we discuss a key concept of modern statistical physics and stochastic dynamics – the fluctuation-dissipation theorem (FDT). In the second part, we introduce the reader to the concepts of molecular simulations and our hydrophobic key-lock model. We motivate our model by the previous studies from Setny and co-workers [48–50, 54, 55]. In particular, we reproduce their results from Ref. [55] for which we also lay out the previous discussion of their most novel findings of the influence of hydrophobicity on key-lock kinetics. Hence, we present the primary analysis tools for the MD simulations in the application to the actual key-lock model. In a third part, we present the theoretical approach of analyzing first passage times. On one hand, we depict the fundamental framework of the first passage time problem. On the other hand, we also elaborate on the previous analysis and interpretations in the example of the hydrophobic key-lock model.

2.1 Fluctuation-dissipation theorem

The FDT relates the relaxation or the response of a dynamic variable to a system's equilibrium fluctuations. It can either be employed to determine the susceptibility from the associated equilibrium fluctuations or to derive an observable's intrinsic fluctuations from its susceptibility.[1]

For the explanations in this chapter, we use the archetypical example of a particle suspended in a solvent, which directly compares to ligand diffusion in water. In an original study, Brown observed that small pollen particles suspended in water undergo an irregular (random) motion [108]. The basic interpretation of the *Brownian* motion was that the movement of the pollen is a consequence of random collisions with the surrounding solvent molecules. In 1905 [109], Einstein formalized this idea and found the resulting first example of the FDT. He neglected the particle's

[1] A susceptibility quantifies the change of an extensive property upon the variation of an intensive property.

inertia in his description, which today is known as *overdamped* Brownian motion. Langevin elaborated on the same type of motion including the inertia term, which is called *underdamped* Brownian motion [110].

2.1.1 Generalized Langevin equation

To show how Brownian motion formally fulfills the FDT, we take a look at the model from Langevin. He formulated a classical, phenomenological equation of motion with two forces that act on the Brownian particle of unit mass m [110, 111]. In the most general formulation, the so-called *generalized* Langevin equation (GLE) is an integro-differential equation

$$m\ddot{q} = -m\int_{-\infty}^{t} \xi(t-t')\dot{q}(t')dt' + \mathcal{F}(t)\,. \tag{2.1}$$

The first term on the r.h.s. $-m\xi(t)\dot{q}$ is the viscous friction force, which, in its most general form, incorporates a *memory kernel* defined by the time dependent friction. For physical reasoning, it is convenient to think of the memory kernel as a decreasing function of $|t|$ with a width of τ_c. The second term represents the fluctuating impacts of the fluid molecules formulated as a random force $\mathcal{F}(t)$. This force is most often considered to be a Gaussian random process with zero mean. If we include the time dependent memory kernel, it is also necessary to take a non-vanishing correlation time for the random force into account. Thus, one similarly assumes that the autocorrelation function (ACF) $g(\tau) = \langle \mathcal{F}(t)\mathcal{F}(t+\tau)\rangle$ is a decreasing function of $|t|$ with a finite correlation time τ_c. Given these definitions, we already implicitly applied the FDT. For example, having the same time scale τ_c for the force correlation and the memory is a direct consequence of the FDT.

2.1.2 First and second FDT

There are two fairly general formulations of the FDT, which, nevertheless, are equivalent [111]. The first FDT connects the complex admittance $\mu(\omega)$ with the Fourier-Laplace Transform of the equilibrium velocity ACF such that

$$\mathrm{Re}\,\mu(\omega) = \beta\int_{-\infty}^{\infty} \langle v(t)v(t+\tau)\rangle e^{i\omega\tau}d\tau \tag{2.2}$$

where $\beta = (k_B T)^{-1}$ denotes the inverse of the thermal energy. The second FDT relates the friction $\xi(\omega)$ (in the frequency domain) with the random force ACF such that

$$\mathrm{Re}\,\xi(\omega) = \frac{\beta}{2m}\int_{-\infty}^{\infty} \langle \mathcal{F}(t)\mathcal{F}(t+\tau)\rangle e^{i\omega\tau}d\tau\,. \tag{2.3}$$

From this second expression, we can derive by inverse Fourier transformation the aforementioned relation that the memory kernel and the random force ACF are linearly related by

$$\xi(\tau) = \frac{\beta}{m} g(\tau) \tag{2.4}$$

such that both are decreasing functions of time with the same correlation time τ_c [111]. The statement that a random process fulfills fluctuations-dissipation refers to equations of the type of Eq. (2.4). Therefore, in a correctly motivated stochastic model, the strength and time correlation of the fluctuating force must be determined by the model's memory kernel and *vice versa*.

In a very simple case, the force ACF $g(\tau) = 2k_B T m \xi \delta(\tau)$ defines a δ-*correlated* random process, which determines a *memory-less* friction kernel. This is also termed a *Markovian* random process. All processes with non-zero, finite time correlations for the random force and the memory kernel are simply called *non-Markovian*.

2.2 Molecular dynamics simulations

In most parts of this thesis, we employ MD simulations to investigate explicit-water dynamics in contrast to other molecular simulation methods, such as Monte Carlo simulations. The latter one is rather generating time independent trajectories that are good for ensemble averages and observations of static properties. The MD scheme numerically solves Newton's equations of motion for the N-body problem of the respective molecular and atomic coordinates. Hence, MD simulations explicitly enable investigations on dynamic properties, which is the goal of this thesis.

In general, molecular simulations require an expression for the potential energy based on the system's configuration, which depends on coordinates, velocities, quantum spins, etc. [112, 113] On one hand, the physical detail of the potential energy determines the information depth of the simulations. On the other hand, highly resolved physics require higher computational costs and possibly capture superfluous information. Thus, in general, the trade-off between accuracy and numerical feasibility must be considered. From the discussions in the first chapter, it is evident that our investigations on the impact of hydrophobicity in key-lock association require at least (a) the level of information of a water model with fixed point charges and (b) a number of water molecules on the order of $\mathcal{O}(10^3)$ or higher. Therefore, an entirely classical approach of molecular mechanics is sufficient and efficient.[2] A classical MD approach usually employs empirical intra- and intermolecular potentials, which are more intuitively termed as bonded $V_{\text{bonded}}(\mathbf{R})$ and non-bonded

[2] Quantum spins can certainly be neglected.

interactions $V_{\text{non-bonded}}(\mathbf{R})$, receptively. Both usually depend on the n atomic coordinates $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_n\}$ and should preferably be analytically expressed (for the numerical treatment).

The bonded interactions define the molecular structure due to chemical bonds. The chemical bonds are expressed in terms of the bond distances $R_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$ between two atoms i and j, in terms of angles α_{ijk} between bond pairs of three atoms i, j and k, as well as in terms of dihedral angles β_{ijkl} between two planes spanned by a set of three bonds with four atoms i, j, k and l.

$$V_{\text{bonded}} = \sum_{i,j}^{\text{bonds}} \frac{K_{ij}^{\text{bond}}}{2}(R_{ij} - R_{ij}^{\text{eq}})^2 \tag{2.5a}$$

$$+ \sum_{i,j,k}^{\text{pairs}} \frac{K_{ijk}^{\text{angle}}}{2}(\alpha_{ijk} - \alpha_{ijk}^{\text{eq}})^2 \tag{2.5b}$$

$$+ \sum_{i,j,k,l}^{\text{dihedrals}} K_{ijkl}^{\text{dihedral}}(1 + \cos(m_{ijkl} \cdot \beta_{ijkl} - \beta_{ijkl}^{\text{eq}})) \tag{2.5c}$$

The constraint to an equilibrium bond distance R_{ij}^{eq} is often expressed by a harmonic spring potential with spring constant K_{ij}^{bond} as it is written in Eq. (2.5a). The harmonic nature of the potential accounts for the minimum of an underlying realistic potential, such as the Morse potential, which is sufficiently approximated up to a second order perturbation. Hence, we assume that high excitations leading to chemical bond breakage, i.e., dissociation, can be neglected. Similarly, a harmonic potential can be used to restrain the angle between two neighboring bonds to its equilibrium angle α_{ijk}^{eq} with spring constant K_{ijk}^{angle} as given in Eq. (2.5b). Last, the potentials used for constraining a molecule's equilibrium dihedral angles β_{ijkl}^{eq} are most often expressed in terms of trigonometric functions with force constants $K_{ijkl}^{\text{dihedral}}$ as shown in the example of Eq. (2.5c). The choice of trigonometric functions in combination with the multiplicity m_{ijkl} enables multiple stable dihedral configurations. If the dihedral angle takes only one stable value, namely if its multiplicity m_{ijkl} is one, the trigonometric function might also be replaced by a harmonic constraint.

If the flexibility of the bond length or angle is not required, they can also be held rigid with the LINCS [114], SHAKE [115], or SETTLE algorithm [116]. The latter is most often used in rigid water models because the chemical bonds between the oxygen and the hydrogen atoms require a high spring constant. These high spring constants would yield high oscillation frequencies, which would have to be resolved by relatively small time steps. Hence, we avoid the resulting superfluous time resolution by applying the SETTLE algorithm for water molecules.

The non-bonded interactions are usually based on classical potentials such as electrostatic interactions and empirical forms of the Pauli repulsion as well as the van-der-Waals attraction. Together the two are most often summarized in the Lennard-Jones (LJ) potential.

$$V_{\text{non-bonded}} = \sum_{i,j}^{\text{Coulomb}} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{2.6a}$$

$$+ \sum_{i,j}^{\text{LJ}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{|\mathbf{R}_i - \mathbf{R}_j|} \right)^{12} - \left(\frac{\sigma_{ij}}{|\mathbf{R}_i - \mathbf{R}_j|} \right)^{6} \right] \tag{2.6b}$$

The electrostatic interactions in Eq. (2.6a) are expressed in terms of the Coulomb potential where q_i and q_j are point charges, and ϵ_0 denotes the vacuum permittivity. The Pauli repulsion is motivated by the quantum mechanical principle that electrons (fermions) in the same state defined by orbital quantum numbers must not spatially overlap. Hence, a steeply diverging energy often models this repulsive contribution by the positive term scaling with r^{-12} in Eq. (2.6b). The negative term, scaling with r^{-6} in Eq. (2.6b), implements the dispersion or van der Waals (vdW) attraction which originates from spontaneous and fast charge fluctuations that induce dipole-dipole interactions between separate atoms.

Throughout this study, we perform MD simulations with the GROMACS-4.6.3 [117] or GROMACS-5.1.2 [118] simulation packages. The first MD studies in Chapter 4 are conducted with the older version (4.6.3) whereas we update to the newer package (5.1.2) for most recent results in Chapter 5.

2.2.1 Water models

Bernal and Fowler described the first empiric water model in 1933 [119]. This model comprised an empiric repulsion, a van der Waals attraction, and three point charges, which were distributed on the hydrogen and oxygen atoms. Equivalent representations were used in 1969 by Barker and Watts in the first Monte Carlo simulations [120] and 1971 by Rahman and Stillinger in the first MD simulations of water [121]. Since these first water simulations, great effort was put into developing improved water models that are suitable for efficient but accurate simulations.

Ouyang and Bettens recently reviewed the water models from the earliest ones to the state-of-the-art in 2015 and grouped them into three classes [122]. The first class comprises the empirical water models such as the one developed by Bernal and Fowlers and similar models, which were employed in the first water simulations mentioned above. The second class contains water models, which include polarizability since the first of its kind in 1992 [123]. The third class groups

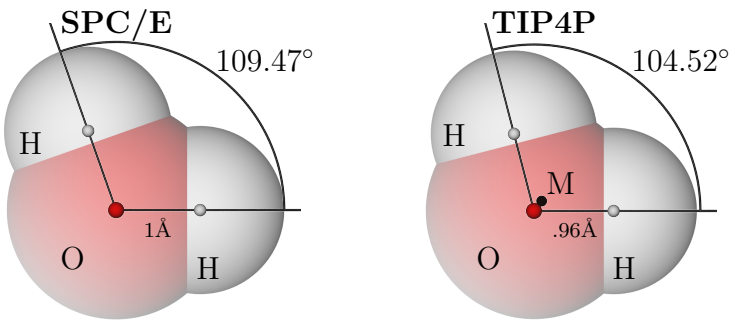

Fig. 2.1: The SPC/E and TIP4P water models have comparable geometries. The bond length is 1 Å in SPC/E and 0.96 Å in TIP4P. The angle spanned by the two hydrogen atoms is 109.47° in SPC/E and 104.52° in TIP4P. The shaded spheres represent an effective soft core size of the atoms. The small spheres show the location of the interaction sites for the Lennard Jones potential (O) and the electrostatic interactions (O and H). In the TIP4P a special massless interaction site (M) carries the negative charge close to the oxygen atom while the oxygen interaction site (O) is uncharged.

the youngest models that use ab-initio data since the first of this kind in 2006 [124, 125]. Within each class, further advances usually rescaled the input parameters to optimize the simulated observables to the experimental ones like density, phase transitions, heat of evaporation, and more [122, 126, 127]. Ouyang and Bettens sorted the parameter revisions into different families [122]. The empirical water class contains the SPC and the TIPnP families from which we illustrate two prominent examples in Fig. 2.1: the SPC/E [126] and the TIP4P model [127]. The latter model is of particular importance for the work presented in this thesis because this model reaches the appropriate accuracy with fast computational speed.

Both models are similar concerning their geometries, which minorly deviate in the exact bond lengths and angles and their LJ interactions. The SPC/E model has a bond length of one ångström while for TIP4P the bond length is 4% smaller, namely 0.96 Å. The angle between the two bonds in SPC/E water spans 109.47° and 104.52° in TIP4P. The bond angle in TIP4P is also roughly $\lesssim$ 4% smaller than in SPC/E water. The geometries are held rigid using the SETTLE algorithm [116], which we previously mentioned. In both models, the respective oxygen position acts as the interaction site for the only LJ potential of the molecules. The potential parameters for SPC/E read $\epsilon_w = 0.650$ kJ mol^{-1} and $\sigma_w = 3.166$ Å, and for TIP4P one uses $\epsilon_w = 0.648$ kJ mol^{-1} and $\sigma_w = 3.154$ Å. Thus, the geometry and LJ interaction are considerably similar in both models. The striking difference between SPC/E and TIP4P lies in the treatment of water's polarity. The SPC/E model carries the partial charges on the atomic sites while the hydrogens (H) are positively charged by 0.4238 e and the oxygen (O) atom carries a negative charge of $-0.8476\,e$. TIP4P includes no charge on the oxygen atom but incorporates an extra, massless virtual

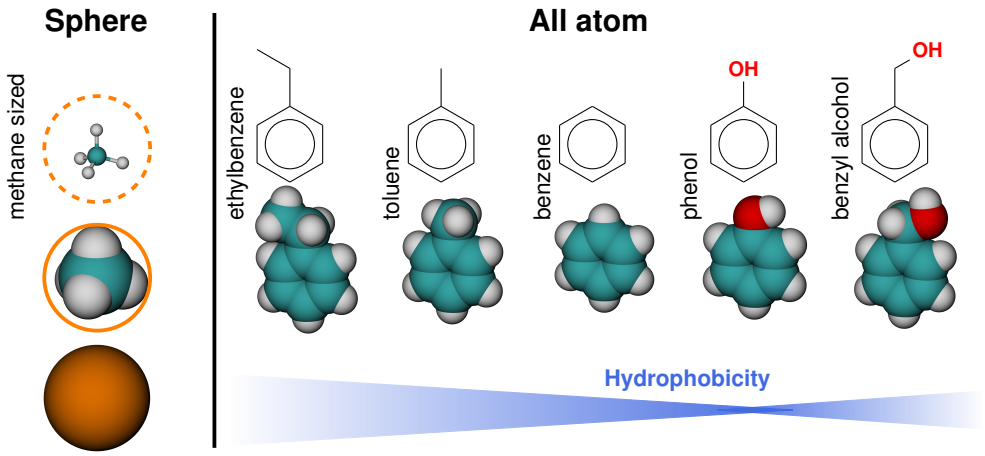

Fig. 2.2: The l.h.s. shows how the spherical LJ particle scales with the size of methane. The radius of the sphere, namely the parameter $\sigma_s = 3.73$ Å, is the same as for the united-atom representation of methane. The r.h.s. shows five aromatic compounds. In principle, the central benzene ring is extended by one more methyl group while stepping to the left to toluene and ethylbenzene. Stepping from benzene to the right, we add a hydroxyl group to create phenol. Farther right we add a hydroxymethyl group to the aromatic ring to form benzyl alcohol. Phenol is the most hydrophilic compound while hydrophobicity increases stepping to the left and right of it. The blue bar at the bottom of the picture illustrates this relative hydrophobicity of all compounds.

interaction site (M) with -1.04 e. This additional interaction site is located 0.15 Å away from the oxygen atom in the angular center between the hydrogen atoms. In Fig. 2.1 the virtual site is represented as a black sphere. Each hydrogen atom carries a positive point charge of 0.52 e accounting for the negative charge on the virtual site.

2.2.2 Organic solutes

Throughout we utilize various hydrophobic compounds as ligands including a spherical solute of the size of methane and all-atom (AA) represented aromatic compounds shown in Fig. 2.2. First, we employ a spherical ligand, which purely interacts with a single LJ potential. In order to mimic a hydrophobic sphere of the size of methane, we use $\sigma_s = 3.73$ Å from the united atom (UA) representation of methane (l.h.s. in Fig. 2.2) [128]. The dispersion interaction strength is $\epsilon_s = 294$ J mol^{-1}, which is smaller than in UA methane. Thus, our spherical particle is even more hydrophobic than the UA methane, which was originally used in Refs [55, 128].

Secondly, we utilize AA representations of the aromatic compounds such as benzene, toluene, ethylbenzene, phenol, and benzyl alcohol as shown on the r.h.s. of Fig. 2.2. All of these employ the LINCS algorithm [114] to constrain the bond lengths while the harmonic angles and improper dihedral angles are treated by

the respective potentials.[3] In general, all (bonded and the non-bonded) potential parameters are taken from the OPLS-AA force field [129]. The carbon atoms interact via a LJ potential as well as negative partial charges. The hydrogen atoms are bound to the carbon atoms and thus carry the opposite positive partial charge such that the molecules are overall electroneutral. The additional oxygen atom on phenol and benzyl alcohol also interact via a LJ potential as well as a negative partial charge. In these cases, the associated hydrogens are also oppositely charged. The partial charges of the hydroxyl group are considerably large such that it can form hydrogen bonds with water. Hence, phenol can be considered to be the most hydrophilic compound while hydrophobicity increases left and right to phenol as represented by the blue bar in Fig. 2.2.

2.2.3 Hydrophobic key-lock model

Our principle MD setup is akin to the hydrophobic key-lock model from publications by Setny [48–50] who investigated the binding of a spherical hydrophobic ligand to a hydrophobic pocket. A reproduction of his system is illustrated in Fig. 2.3. The ligand is constrained to the symmetry axis of the pocket, which is illustrated as z-axis. In particular, the ligand's center of mass is restrained by external harmonic potentials in x- and y-direction with spring constants $k_x = k_y = 42000$ kJ mol^{-1} nm^{-2}. The spring constants are chosen accordingly to Ref. [55]. Therefore, all our investigations treat one-dimensional ligand motion along the z-direction. The pocket is molded into the gray wall, which is essentially a crystal of LJ spheres. The first planar crystal layer that is in contact with water defines the position $z = 0$ as indicated in Fig. 2.3. The actual simulation box contains two walls that face each other with a slab of 1545 TIP4P water molecules in between. This water slab has a thickness of roughly ~ 4 nm. Together, both pockets effectively correspond to a single binding site with a reflective boundary at a distance of $z_{max} \sim 2$ nm. The box size in z-direction extends up to 10 nm such that the volume behind the pockets is empty. The box lengths in x- and y-direction are 3.375 nm and 3.464 nm, respectively. The walls terminate at the periodic boundaries in x- and y-direction such that the crystal structure continues without gap into the periodic images.

In particular, we construct a versatile pocket, which can be adjusted in curvature and depth. The wall itself is composed of pseudo-atoms that are arranged in a hexagonal close packed (HCP) crystal structure, which is illustrated on the l.h.s. of Fig. 2.3. The lattice constant a in the hexagonal plane is $a = 1.25$ Å. A second

[3] In contrast to the empiric water models, we need to use LINCS because it applies the method of Lagrange multipliers to constrain molecules with arbitrary numbers of atoms [114]. SETTLE exactly/analytically solves the underlying equations for three atoms, which is perfect for water models [116].

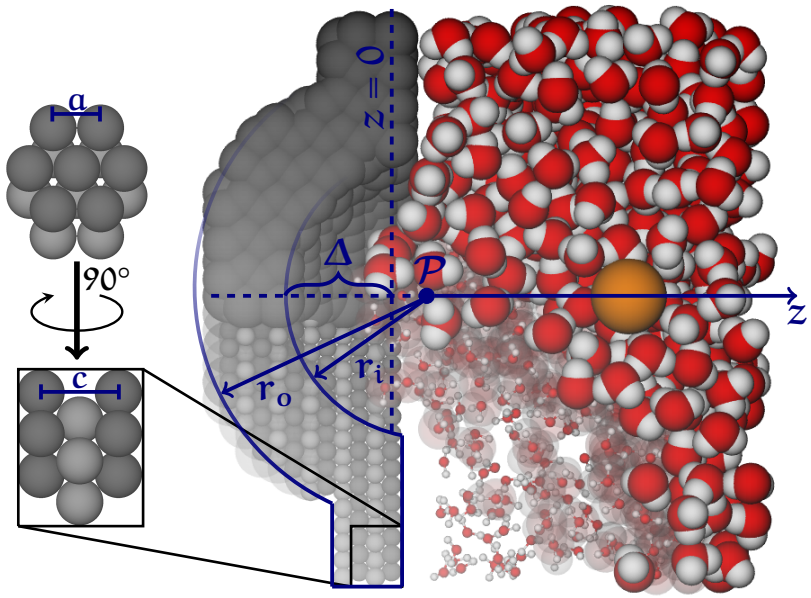

Fig. 2.3: The hydrophobic key-lock model comprises a hydrophobic wall (gray) with a concavely shaped pocket as the binding site, TIP4P water (red) and a spherical ligand (orange). The hydrophobic wall is composed of a HCP crystal structure with ABAB stacking of hydrophobic particles (gray). Light gray and dark gray spheres only differ in color to illustrate the stacking. The lattice constant of the hexagons is $a = 1.25$ Å and the distance between every other stack (A to A, or B to B) is $c = \sqrt{8/3}a \approx 1.02$ Å. Two spherical probe volumes with radii r_i and $r_o = r_i + 5$ Å are centered at point $\mathcal{P}$ and determine the pocket's geometry. By shifting $\mathcal{P}$ or changing the radius r_i the pocket size can be tuned regarding depth and curvature. A given geometry is uniquely set by the tuple (r_i, Δ), where $\Delta = r_i - \mathcal{P}$ is the pocket depth.

hexagonal layer stacks into the depressions of the first one such that the HCP crystal structure arranges in an ABAB stacking. The orthogonal distance of neighboring layers (AB) is $0.5c = 0.5\sqrt{8/3}a \approx 1.02$ Å. The lattice constants a and c are relatively small compared to realistic atomic crystal structures because we want to design small and finely grained pockets. The LJ parameter for one of the wall's pseudo-atoms read ($\sigma_p = 4.15$ Å, $\epsilon_p = 2.40$ J mol^{-1}) whereas the dispersion attraction is also chosen relatively small compared to realistic atoms because of the high crystal density. Further, the dispersion attraction is even weaker than in Refs [48–50, 55], which we further describe in Appendix A. During the simulations, the wall particles are simply *frozen*, namely, their positions are fixed and are not updated during the numerical integration. Thus, we do not make use of bonded interactions for a possible flexibility of the wall.

In Chapter 4, we construct various pockets with varying geometries. The geometry of these pockets is determined by the parameters $\mathcal{P}$, r_i and r_o (see also Fig. 2.3). The point $\mathcal{P}$ is the center of two concentric spherical volumes one with radius r_i, defining the inner radius of the pocket, and another one with radius r_o, defining the

outer larger radius of the back of the pocket. The distance d_j of a pocket atom j to the point $\mathcal{P}$ fulfills the condition $r_i < d_j < r_o$. We keep the pocket wall thickness constant such that r_o is determined by r_i, i.e., $r_o = r_i + 5$ Å. Additionally, the number of layers at the pocket bottom as well as at the wall's outer edges is limited to four. In summary, if the inner radius r_i is tuned the curvature of the pocket changes. If we shift the position $\mathcal{P}$ along the z-axis, the pocket is deepened or flattened. Our particular choices for the pocket geometry are summarized at the proper place in Chapter 4. In the following sections, we exemplify the methodology utilizing a pocket for which $r_i = 9.5$ Å and $\mathcal{P}$ lies at $z = 1.5$ Å. We define the pocket depth $\Delta = r_i - 1.5$ Å $= 7.0$ Å. The tuple of these parameters (r_i, Δ) fixes a given pocket geometry. We consider the (9.5, 7)Å geometry as our reference pocket throughout the thesis.

In general, all simulations are prepared as *NVT* ensembles at $T = 298$ K with a Nosé Hoover thermostat. The thermostat's coupling frequency is 1 ps during equilibration runs and 10 ps during production runs. Note that only the water is coupled to the thermostat throughout the production runs. Although for technical reasons, we use constant volume simulations, the density of the water slab is tuned to $\rho = (997 \pm 0.3)$ g/dm^3 by slight adjustments of the mirrored walls' separation.

Next, we explain the analysis tools for MD simulations; however, we already show and apply them to our simulation results of one given pocket, our reference pocket. The discussion and interpretation follow the previous findings from Setny *et al.* [55] and basically reproduce their results. This way we can introduce the reader to the details of the particular kinetics in hydrophobic key-lock binding.

2.2.4 Constrained sampling

We utilize the umbrella sampling scheme to probe how the potential, the ligand friction, and the pocket hydration fluctuations depend on the ligand-pocket separation. Also, previous related studies [55, 101] applied this scheme.

In a single umbrella simulation, the ligand is restrained to a given position z_i^r using an external harmonic potential

$$u_i(z) = \frac{k_z}{2}(z - z_i^r)^2 \tag{2.7}$$

where we use the value for the force constant to be $k_z = 835$ kJ mol^{-1} nm^{-2}. Our first so-called *umbrella window* constrains the ligand inside the pocket, i.e., to a position that is 3 Å away from the pocket's bottom. Thus, in the example of the

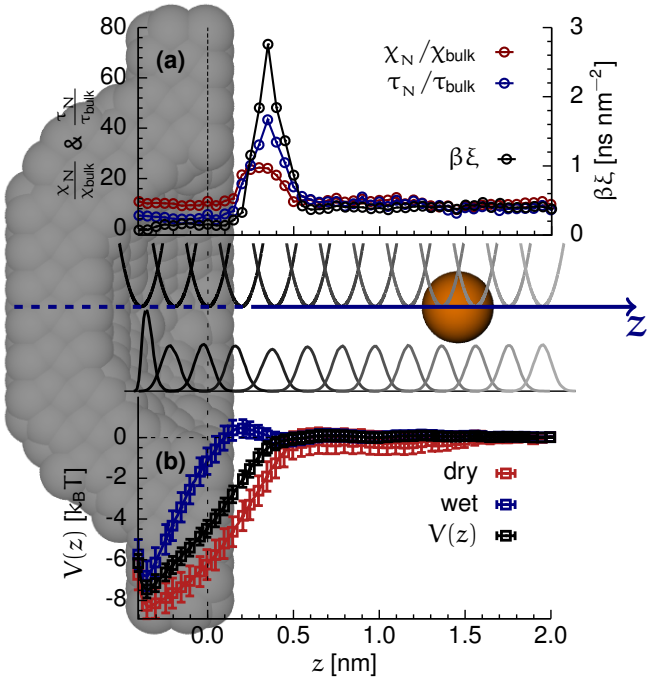

Fig. 2.4: The umbrella sampling simulations resolve the PMF and kinetic measures, such as ligand friction and solvent fluctuations. (a) The ligand friction ξ (black circles), the pocket water occupancy fluctuations χ_N (red circles), and their time scale τ_N (blue circles) maximize if the ligand is in front of the pocket. (b) The PMF is steeply attracting whereas this slope shifts with the pocket hydration state. For wet pocket states ($N > 5$) the attracting slope lies inside the pocket, and a small barrier arises prior to binding (blue squares). For the dry pocket state ($N = 0$) the attracting slope reaches far outside the pocket (red squares). The average PMF, disregarding the pocket hydration, lies in between these two limits (black squares).

(9.5, 7)Å geometry, we have $z_1^r = -4$ Å.[4] We choose a total number of $S = 49$ umbrella windows at increasingly z-values with a spacing of $\Delta z = z_{i+1}^r - z_i^r = 0.5$ Å. (In the middle of Fig. 2.4 every other umbrella potential is drawn on top of the blue z-axis. The color transitions from black to gray while the restraining position increases.) For each umbrella window, the ligand trajectory can be sorted into an umbrella-like shaped position histogram. (The histograms are drawn below the blue z-axis in Fig. 2.4.) For subsequent evaluation of the potential of mean force (PMF), the histograms must overlap. Therefore, we choose a sufficiently small spacing and the spring constant accordingly to Ref. [55].

We calculate the PMF using the weighted histogram analysis method (WHAM) [130, 131], which requires the umbrella-like shaped histograms of the ligand. In particular, for each window $i = 1, \ldots, S$, the ligand position is binned into a histogram with n_{il} counts in bin l centered around z_l ($l = 1, \ldots, M$). These histograms also define $N_i = \sum_{l=1}^{M} n_{il}$ as the number of samples in the umbrella window at z_i^r. The objective

[4] In principle, the value of the first restraining position z_1^r varies with pocket depth Δ, which is why we state that the first umbrella potential is 3 Å away from a pocket's bottom.

is to construct the underlying unconstrained free energy or PMF, which is defined by the unconstrained equilibrium probability distribution p_l via

$$V(z_l) = -k_B T \ln(p_l/\delta_l) \tag{2.8}$$

where δ_l is the bin width of the histograms. (We use $\delta_l = 0.1$ Å.) The probability distribution is obtained by iteratively solving the following equations

$$f_i = \frac{1}{\sum_l c_{il} p_l} \tag{2.9a}$$

$$p_l = \frac{M_l}{\sum_i N_i f_i c_{il}} \tag{2.9b}$$

where $c_{il} = \exp(-\beta u_i(z_l))$ is the Boltzmann factor of the biasing umbrella potential from Eq. (2.7), and $M_l = \sum_i n_{il}$ is the total count in the l-th bin from all umbrella windows. The first iteration can initiate from f_i in Eq. (2.9a) taking, for instance, an arbitrary $p_l = 1$ for every l. Finally, for error estimation, we choose the variance calculations introduced by Zhu and Hummer [131] via

$$Var_j(V(z)) = (\beta k_z \Delta z)^2 \left[\frac{Var_1(z) + Var_j(z)}{4} + \sum_{i=2}^{j-1} Var_i(z) \right] \tag{2.10}$$

which is based on the variance $Var_i(z)$ of the position distribution in each umbrella window. One can then assume that $Var(V(z_1)) = 0$ such that the error is basically accumulating along the z-coordinate.

The black squares with error bars in Fig. 2.4 (b) represent the average PMF $V(z)$ of the spherical ligand upon binding to our reference pocket. The pocket is strongly attracting the ligand for $z < 4$ Å. From Chapter 1 and Refs [52, 55], we also know that the water in such hydrophobic pockets can evaporate such that the pocket can be temporarily dry or wet. These two different states impact the mean force on the ligand. For instance, if we sample the histograms for WHAM for given hydration states of the pocket, we observe different PMFs. If we sample only dry states, namely when zero water molecules ($N = 0$) are inside the pocket, the attracting slope of the PMF is shifted further away from the wall (red squares). In contrast to the dry PMF, if we sample wet states for which the number of water molecules N inside the pocket exceeds five, the attracting slope shifts closer to the wall and even exhibits a small barrier (blue squares). Thus, the different hydration states yield different temporal energy landscapes that are spatially shifted to one another. Such findings were already reported in Ref. [101].

The ligand friction maximizes in front of the pocket, prior to binding, which is plotted as black circles in Fig. 2.4 (a). This local friction $\xi(z_i^r)$ of the ligand is also obtained from the umbrella simulations by

$$\beta\xi(z_i^r) = \frac{\int_0^\infty \langle \delta z(t)\delta z(0)\rangle_i dt}{\langle \delta z^2\rangle_i^2} \tag{2.11}$$

where $\langle \delta z(t)\delta z(0)\rangle_i$ is the position ACF (PACF) with $\delta z(t) = z(t) - \langle z\rangle_i$ [132]. At the end of this chapter, we see that the friction obtained from Eq. (2.11) cannot be reproduced by a second, Markovian method. This issue of Markovianity and non-Markovianity is one focus of the thesis and accompanies many discussions throughout. It was also previously highlighted in Ref. [55] whereas the long-time correlations in position and force correlations were presented.

Regarding the pocket hydration itself, we also utilize the umbrella trajectories to calculate the average pocket water occupancy $\langle N\rangle$ and its temporal ACF

$$C_{\delta N\delta N}(t|z_i^r) = \frac{\langle \delta N(t)\delta N(0)\rangle_i}{\langle N\rangle_i} \approx \chi_N(z_i^r)e^{-t/\tau_N(z_i^r)} \tag{2.12}$$

where $\delta N = \langle N(t) - \langle N\rangle_i\rangle_i$ with $\langle \ldots\rangle_i$ denoting the average taken in the umbrella window at z_i^r. The approximate relation on the r.h.s of Eq. (2.12) defines the fluctuation strength $\chi_N(z)$ and its time scale $\tau_N(z)$. Throughout this thesis, we absorb the explicit spatial dependence into the indices of the observables, e.g., $\chi_N(z) = \chi_{N,z}$. The fluctuation strength and its time scale are two important measures that characterize the hydrophobicity of a surface. For instance, the fluctuation time scale $\tau_{N,z}$ quantifies slow (capillary) fluctuation modes. We fit the exponential tail of the simulation sampled ACF to obtain the fluctuation time scale. More importantly, the water fluctuation strength χ_N has been recently established as an index for a surface's degree of hydrophobicity [133–140]. In the thermodynamic limit (i.e., for large N), the fluctuation strength relates to the isothermal compressibility [133]. The fluctuation strength is proportional to the second cumulant of the pocket water occupancy distribution of a chosen probe volume (here the pocket), which in turn quantifies the free energy of desolvating the volume of interest. In particular, large values of χ_N indicate tendencies for surface dewetting and, thus, express enhanced hydrophobicity/hydrophobic confinement.

Throughout the thesis, we compare $\chi_{N,z}$ to the bulk value χ_{bulk} quantifying the fluctuation strength in a comparably sized bulk probe volume far away from the wall. Thus, the ratio $\chi_{N,z}/\chi_{bulk}$ gives a measure of how hydrophobic the pocket is if compared to the likelihood of cavity formation in the bulk water slab. The

fluctuations in Fig. 2.4 (a) for large ligand-pocket separations ($\lim_{z\to\infty}\chi_N(z) = \chi_{N,\infty}$) are already ten times higher inside the pocket than in the bulk. While the ligand is restrained at smaller z values, the water fluctuations in the pocket peak around $z = 3.5$ Å. Similarly, the fluctuation time scale τ_N is enlarged compared to the bulk measure τ_{bulk} and maximizes before ligand binding.

2.2.5 Unconstrained sampling

For a given key-lock setup (with a given pocket geometry and a given ligand), we prepare over a thousand free binding trajectories in which the ligand is free to move in z-direction. First of all, we store a preliminary production run of 20 ns in steps of 0.2 ps in which the ligand is fixed/frozen at the reflective boundary. This initial trajectory serves as a source for randomly seeded starting configurations for subsequent binding simulations. We randomly pick a frame from the preliminary simulation run and extend it by a further annealing procedure to ensure the initial configuration's randomness. Within a short simulation of 50 ps, the system is heated up to 350 K using a stochastic integrator scheme, after which, the heated configuration is equilibrated for another 100 ps to 298 K using the Velocity Verlet algorithm. In the final production run, the ligand is free to move in the z-direction. This simulation is terminated once the ligand is bound two ångströms within the pocket, namely at $z < z_f = -2$ Å. The time for binding, i.e., the first passage time (FPT), is then sorted to calculate the mean FPT (MFPT), $\mathcal{T}(z)$, depending on position z. The blue error bars in Fig. 2.5 represent the simulation sampled MFPT curve of the spherical ligand to bind into our reference pocket geometry.

2.3 First passage time problems

In general, the problem of the FPT for an arbitrary random process is solved if the distribution of the time to cross a given boundary is known. Often, the primary goal is the first moment of the FPT distribution, i.e., the MFPT. The knowledge of the higher moments, however, are useful to formulate the probability generating function, which determines the FPT distribution and, hence, the full problem.

Formally, the moments $\mathcal{T}_n$ of the FPT distribution can be derived via the *adjoint operator approach*, which was generalized to non-Markovian processes by Hänggi and Talkner [141, 142]. This approach requires the construction of the adjoint operator $\Omega^\dagger$, which is determined by the evolution equation.[5] This operator determines a

[5] For Gaussian white noise the Fokker-Planck equation is the evolution equation.

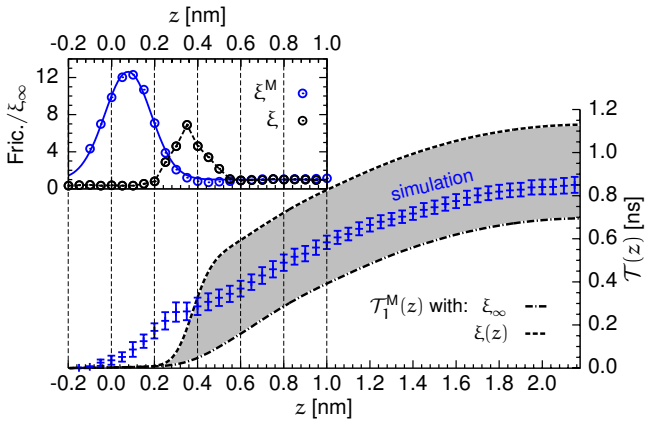

Fig. 2.5: The MFPT $\mathcal{T}_1^{\mathrm{M}}(z)$ from Eq. (2.14) taking into account a constant friction (dash dotted line) or the friction from PACFs in Eq. (2.11) (dashed line) respectively under- or overestimates the results from simulation (blue error bars). The gray shaded area marks an error range that opens due to the predictions which serve as lower and upper boundary. The inset compares the friction profile from PACF $\xi(z)$ from Eq. (2.11) (black circles) and the friction profile $\xi^{\mathrm{M}}(z)$ from Eq. (2.15) (blue circles).

recursive equation for the moments of the FPT distribution, which is formally written as

$$\Omega^{\dagger}\mathcal{T}_n = -n\mathcal{T}_{n-1}\,. \tag{2.13}$$

For Markovian processes, the adjoint operator approach is straightforward and exact [143]. The solution yields an expression that incorporates the free energy landscape $V(z)$ and a possibly spatially dependent friction $\xi(z)$

$$\mathcal{T}_n^{\mathrm{M}}(z, z_f) = n\beta \int_{z_f}^{z} dz' \xi(z') e^{\beta V(z')} \int_{z'}^{z_{max}} dz'' e^{-\beta V(z'')} \mathcal{T}_{n-1}^{\mathrm{M}} \tag{2.14}$$

where the zeroth moment $\mathcal{T}_0^{\mathrm{M}} = 1$ determines normalization of the FPT distribution. The reflective z_{max} and absorbing boundary z_f define the one-dimensional spatial interval from which the escape or first passage process takes place [144, 145]. For $n = 1$, Eq. (2.14) is exactly the MFPT which Setny *et al.* [55] employed for theoretical estimates of the binding time for the hydrophobic key-lock binding process.

If using the friction $\xi(z)$ and the PMF $V(z)$ from the previous umbrella sampling in Fig. 2.4, the mean binding time from simulation in Fig. 2.5 (blue error bars) is drastically overestimated (black dashed line). If taking the PMF alone and assuming a constant friction, namely the bulk value $\beta\xi(z) = \beta\xi_{\mathrm{bulk}} = 0.4$ ns nm^{-2}, the mean binding time curve (black dash-dotted line) is underestimating the simulation result. Hence, the Markovian assumption in Eq. (2.14) seems to break down for the hydrophobic key-lock model [55]. Note that $\mathcal{T}(z)$ denotes the simulation sampled

MFPT where $\mathcal{T}_1^{\mathrm{M}}$ refers to the Markovian theory. This notation is kept throughout the whole thesis.

Nevertheless, we can extract a "kinetic" profile that illuminates the dissipation, i.e., the effective deceleration in the simulation sampled binding trajectories. Upon solving the case $n = 1$ of Eq. (2.14) for the spatially dependent friction [146, 147] one obtains

$$\beta\xi^{\mathrm{M}}(z) = \frac{\partial\mathcal{T}_1^{\mathrm{M}}(z)}{\partial z}\frac{e^{-\beta V(z)}}{\int_z^{z_{\max}} dz' e^{\beta V(z')}} \tag{2.15}$$

which differs to the friction that is sampled from umbrella simulations. This discrepancy becomes evident in the inset of Fig. 2.5 where the result from Eq. (2.15) (blue circles) peaks larger and is spatially shifted in comparison to the friction from umbrella simulations and Eq. (2.11) (black circles). However, the result from Eq. (2.15) naturally reproduces the simulation sampled MFPT via Eq. (2.14) by construction.

3 The statistical mechanics of hydrophobic key-lock kinetics

In this chapter, we model the binding of a ligand to a hydrophobic pocket that undergoes bimodal wet-dry transitions using a minimalistic stochastic model which comprises two particles in one dimension. One particle is the ligand and a second pseudo-particle models a bimodally fluctuating water interface in the pocket. Both particles couple via an interaction that is motivated by solvation free energy. Hence, the model captures two fundamental ingredients of hydrophobic key-lock binding: bimodal hydration fluctuations of the pocket and a strong attraction that couples the ligand and the water interface due to hydrophobic solvation. Mathematically, we formulate the two particle system as two coupled stochastic equations. In Section 3.1, the formal details of the model are described. Section 3.2 presents numerical calculations of the model while we focus on the binding kinetics. A comparison of mean binding times to a corresponding memoryless stochastic process demonstrates the breakdown of Markovian behavior if only the ligand distance serves as reaction coordinate. A maximized friction before binding indicates that additional damping in hydrophobic key-lock association originates from a strongly fluctuating coupling of the water interface and the ligand. Our analysis closely follows the procedure of the MD setup described in the previous chapter but gives additional physical insight for the origin and nature of the non-Markovian kinetic effects. To further corroborate these findings, Section 3.3 deals with a complementary theory describing the ligand motion alone using a generalized Langevin equation including a memory kernel which models the ligand-water coupling as retarded damping. This approach enables further interpretation and an analytical quantification of the numeric results of the enhanced friction in the two particle model and the MD simulations in subsequent chapters.

3.1 Stochastic model

Fig. 3.1 illustrates the two particle model where the thick blue line represents the water interface at the coordinate z_s, and the orange sphere is the ligand of radius R at z_l. The ligand diffuses with the properties of a spherical particle in water utilizing Stokes friction and the Einstein relation $D = k_B T/6\pi\eta R$, where η denotes the viscosity of the bulk solvent [109]. We choose the same diffusivity constant for

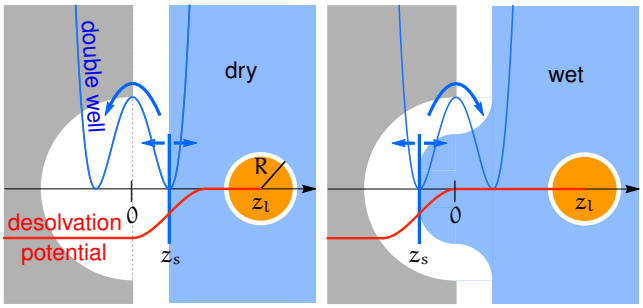

Fig. 3.1: Illustration of the stochastic model: The water interface (vertical blue bar) at position z_s fluctuates inside a double-well potential (blue curve, Eq. (3.1)) leaving the pocket either *dry* (left) or *wet* (right). The ligand (orange sphere) at position z_l diffuses freely in one dimension on the z-axis, which is orthogonal to the wall. It interacts with the interface by an attractive *desolvation* potential (red line, Eq. (3.3)). The interaction potential moves with the interface. Note again that particle motion is restricted to one dimension. The schematics of the wall with hemispherical pocket (gray) and bulk water area (blue) are only shown for illustration.

the water interface pseudo-particle because the kinetic coupling known from the MD simulations occurs if water fluctuations and the ligand's diffusion time scale are comparable, i.e., on the same order of magnitude.

Quantitatively, we scale the system according to to the physical constants of the explicit-water MD simulations in Chapter 2 and Ref [55]. Therefore, the effective temperature during the numerical calculations of our stochastic model relates to $T = 300$ K. The particle size of the ligand is that of a methane molecule, which is roughly $R = 0.4$ nm. Together with the viscosity $\eta \approx 10^{-3}$ Pa·s of water [148], the corresponding diffusion constant relates to $D = k_B T/6\pi\eta R = 0.54\,\mathrm{nm^2\,ns^{-1}}$. Though to enable a general presentation of the results, we scale the thermal energy $k_B T$, the length scale R, and the diffusion constant D to one thus setting them as the natural scale in our stochastic model. This also determines our Brownian time scale $\tau_B = R^2/D$ (≈ 0.3 ns).

Again, the water interface is treated as Brownian pseudo-particle and is subject to an external double-well potential

$$V_{dw}(z_s) = \frac{h}{R^4}\left(z_s^2 - R^2\right)^2 + b \cdot z_s \tag{3.1}$$

drawn as blue curve in Fig. 3.1. If the biasing parameter is $b = 0$, we define that the positions of the two wells are situated at $z = \pm R$, whereas h is the height of the barrier, which lies at $z = 0$. To further enable changes in relative depths of the potential wells, we introduce the bias given by the non-zero linearity constant b in units of $k_B T/R$. The double-well potential implements the dry and wet pocket states in which water resides either in front of the pocket or inside of it.

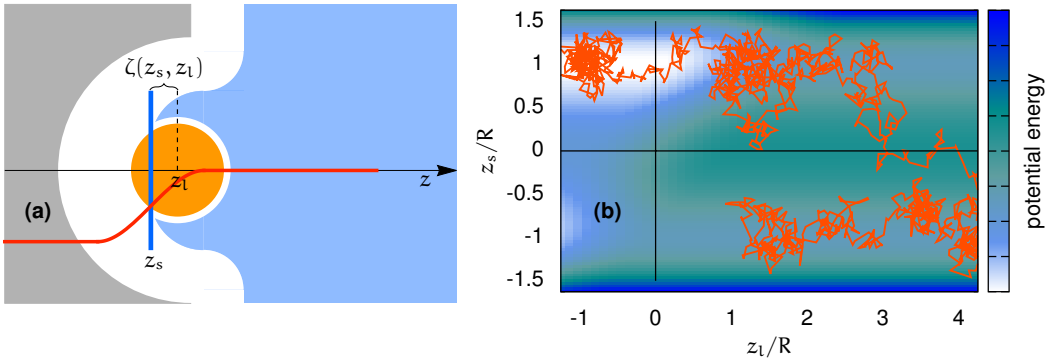

Fig. 3.2: (a) If the distance between the ligand (orange) and the water interface (blue bar) is smaller than the ligand radius, i.e., $|\zeta| = |z_l - z_s| < R$, they interact via the desolvation potential (red curve, Eq. (3.3)). The interaction scales with the solvated volume of the ligand, which is the portion on the r.h.s. of the interface. (b) The stochastic key-lock model essentially describes Brownian motion in two dimensions (z_s, z_l). Here the plot draws an example trajectory (orange) on the potential energy landscape expressed by Eq. (3.5) with arbitrary potential parameters.

The coupling potential, namely the pair potential, between the water interface and the ligand accounts for the free energy gain of desolvation if the hydrophobic ligand passes through the water interface (see Fig. 3.2 (a)). The resulting *desolvation* potential is designed such that it drags the ligand out of the water into the pocket. At the same time, following the principle of action-reaction, the interaction pulls the pseudo-particle out of the pocket, which conceptually corresponds to the ligand induced drying transition in the MD simulations. For small solutes, the solvation free energy approximately scales linearly with the solvent excluded volume $\Delta G \propto Vol$, whereas after a transition at a crossover length-scale l_c it is proportional to the solvent accessible surface area A, $\Delta G = \gamma \cdot A$ with γ as surface tension [7]. Modeling the microscopic key-lock binding with a comparably small sized ligand, we choose the desolvation potential to scale linearly with the solvent-excluded volume. We demand a reasonable proportionality constant Γ to fulfill $\Delta G(l_c) = \Gamma \cdot Vol(l_c) \equiv \gamma \cdot A(l_c)$ at the crossover length-scale, which thus yields

$$\Gamma = 3\gamma/l_c \,. \tag{3.2}$$

For the parametrization of the desolvation potential, we assume the crossover length-scale at $l_c = 1$ nm, which relates to $l_c = 2.5\,R$. With the surface tension $\gamma \approx 15.36\ k_B T\ nm^{-2}$ for water [149], we calculate the desolvation scaling constant to be $\Gamma = 3\gamma/l_c = 46.1\ k_B T/nm^3 = 2.95\ k_B T/R^3$.

Once the water interface intersects the ligand, the solvent excluded volume starts to decrease as a function of ligand-interface distance, $\zeta(z_s, z_l) = z_l - z_s$ leading

to a favorable attraction (see Fig. 3.2 (a)). The desolvation potential can thus be analytically expressed by

$$\Delta G(z_s, z_l) = \Gamma \left[\frac{4\pi}{3} R^3 - \frac{\pi}{3} (R - \zeta)^2 (2R + \zeta) \right] \tag{3.3}$$

which yields a parabolic pair force, acting on particle $i = s, l$ (short for solvent or ligand) via

$$F^i_{sol}(z_s, z_l) = -\frac{d\Delta G(\zeta)}{d\,\zeta} \frac{d\zeta}{dz_i} = \pi\Gamma(\zeta^2 - R^2) \frac{d\zeta}{dz_i} \tag{3.4}$$

which only applies for $|\zeta(z_s, z_l)| < R$. Note that the gray wall in Fig. 3.1 embedding the pocket is only drawn representatively. A steric repulsion and/or a vdW attraction to the wall are omitted. The former is not needed since we employ a fully absorbing boundary condition in the pocket (at $z_f/R = -1.25$). Inclusion of an additional vdW potential between the ligand and the wall just adds an energetic contribution that is constant in time. The vdW energy would only serve as a constant energy shift that would slightly tilt the resulting PMFs. Thus, the general implications of this chapter about ligand kinetics, friction, and memory would be unchanged. Additionally, we employ a reflective boundary at $z_{max}/R = 5$ in order to avoid the ligand to diffuse too far away from the pocket.

In summary, our model describes Brownian motion of two interacting particles, which effectively span the two-dimensional space of z_s and z_l. The sum of the double-well and desolvation potential is written as

$$V_{tot}(z_s, z_l) = \Delta G(z_s, z_l) \cdot \Theta\left[R - |\zeta(z_s, z_l)|\right] + V_{dw}(z_s) \tag{3.5}$$

which makes evident that the coupling simply yields an associated two-dimensional potential energy landscape. An example for arbitrary potential parameters is illustrated in Fig. 3.2 (b) together with an example trajectory. Our non-linearly coupled, overdamped equations

$$\xi_s \dot{z}_s(t) = +\pi\Gamma \left(R^2 - \zeta(z_s, z_l)^2\right) \cdot \Theta\left[R - |\zeta(z_s, z_l)|\right] - 4\frac{h}{R^4} z_s (z_s^2 - R^2) - b + F_s(t) \tag{3.6a}$$

$$\xi_l \dot{z}_l(t) = -\pi\Gamma \left(R^2 - \zeta(z_s, z_l)^2\right) \cdot \Theta\left[R - |\zeta(z_s, z_l)|\right] + F_l(t) \tag{3.6b}$$

describe the key-lock model with ξ_i as friction coefficients, $i = s, l$, and $F_i(t)$ denoting δ-correlated, i.e., Markovian, random forces fulfilling the fluctuation-dissipation theorem such that $\langle F_i(t) F_i(t_0) \rangle = 2\, k_B T \xi_i \delta(t - t_0)$. Note that both $\xi_s = \xi_l$, since we equalize the diffusivities of both ligand and pseudo-particle.

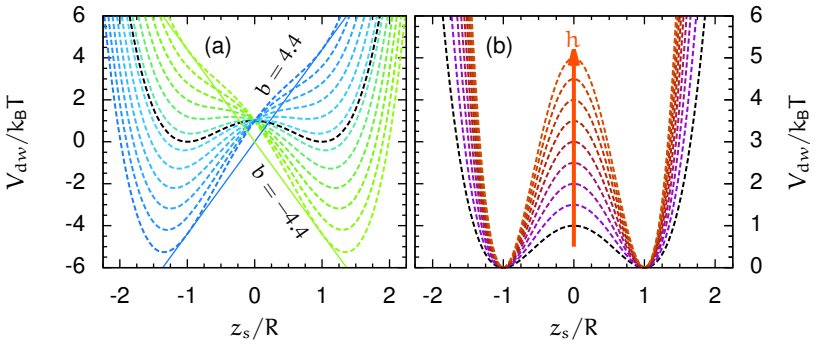

Fig. 3.3: The double-well potential in Eq. (3.1) represents the effective pocket hydration. We tune the potential by (a) biasing from $b = -4.4\ k_BT/R$ to $4.4\ k_BT/R$ in steps of $0.8\ k_BT/R$ or (b) barrier height from $h = 1\ k_BT$ to $5\ k_BT$ in steps of $0.5\ k_BT$. The black dashed lines draw the reference double-well setting ($h = 1\ k_BT$, $b = 0$). Changes in biasing b and magnitude h are indicated by example guide lines in panel (a) or the arrow in panel (b), respectively. The color coding is adopted throughout this chapter.

3.2 Numerical calculations

In general, pocket hydration can be affected by changing the pocket's chemical composition or geometry, which together can be summarized as the degree of hydrophobic confinement. Such changes, however, simultaneously affect the pocket water occupancy and the hydration fluctuation time scale. In our model, we are able to disentangle both effects, occupancy and time scale, and their influence on ligand binding. Here the effective water occupancy can be tuned by changes in the biasing parameter b of the double-well potential from Eq. (3.1). We employ potentials with a biasing ranging from $b = -4.4\ k_BT/R$ to $b = 4.4\ k_BT/R$ in steps of $0.8\ k_BT/R$ as it is shown in Fig. 3.3 (a)). The black dashed line in Fig. 3.3 refers to the reference with barrier height $h = 1\ k_BT$ and zero biasing. Additionally, we separately tune the barrier height from $h = 1\ k_BT$ to $h = 5\ k_BT$ in steps of $0.5\ k_BT$ as it is plotted in Fig. 3.3 (b). Changes in the barrier height directly influence the wet-dry transition time, i.e., the effective evaporation rate. In particular, we know from Kramer's rate theory that the rate of crossing the double-well barrier scales exponentially with barrier height [150]. Note that the color coding from both plots in Fig. 3.3 is consistently adopted to other plots throughout this chapter.

Further, we note that the equilibrium distribution of the water interface depends on the ligand position. The non-linear coupling from the contributions of the desolvation potential in Eq. (3.3) acts in Eq. (3.6), if $|\zeta| < R$. A schematic plot in Fig. 3.4 illustrates how the bimodal distribution changes upon the approach of the ligand. If the ligand and the interface interact, the bimodal distribution of the interface is tilted, which depicts the equilibrium hydration of the pocket. Hence, we observe that the modeled capillary evaporation inside the pocket is enhanced

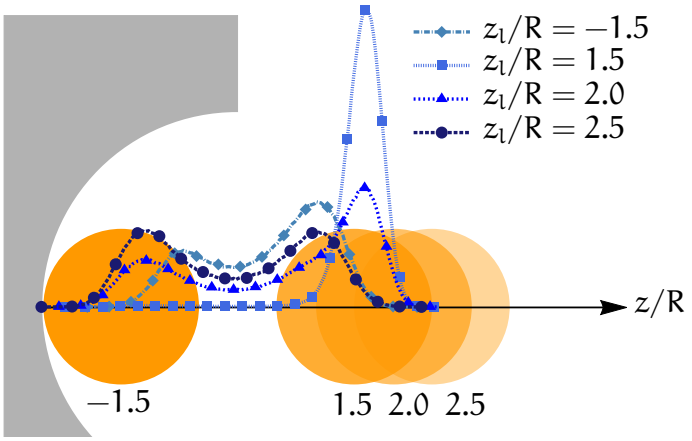

Fig. 3.4: The ligand position affects the equilibrium distribution (blue lines) of the water interface due to the desolvation potential in Eq. (3.3) and thus Eq. (3.5). Hence, the bimodality of the water interface distribution is lost for intermediate states, e.g. $z_l/R = 1.5$, but it is recovered if the ligand is fully bound to the pocket, e.g., $z_l/R = -1.5$.

for close ligand positions, which coarsely mimics how pocket hydration couples to ligand position in all-atom and implicit solvent simulations [52, 55].

We evaluate the binding kinetics of our stochastic model by numerical integration. Therefore, we use the numeric scheme proposed by Ermak and McCammon [151]. We focus on how the coupling affects the ligand kinetics along the z-coordinate. Restricting our observations on the ligand coordinate alone effectively projects the Markovian two-dimensional treatment onto the reaction coordinate z_l. We thus omit the associated subscript such that $z_l \equiv z$ throughout the following Sections 3.2.1 and 3.2.2.

3.2.1 Mean first passage time and memory

As a primary measure of the ligand binding kinetics, we sample the MFPT from each point z to the absorbing boundary at $z_f/R = -1.25$. Therefore, we generate around 2×10^5 trajectories for each setup with given biasing b and barrier height h for which the ligand starts at $z_{max}/R = 5$. The resulting MFPT curves $\mathcal{T}(z, z_f)$ describe the mean binding time of the ligand crossing $z_f/R = -1.25$, given it started at z with a reflective boundary at $z_{max}/R = 5$. In Appendix B.1, we briefly discuss how the results of the MFPT and the dissipative forces change if z_{max}/R takes other values.

Fig. 3.5 (a.1) shows the MFPT curves corresponding to the setups with varying bias of the double-well potential. If the bias is negative (greenish lines), namely if the pocket is preferentially dry, the ligand's mean binding time is faster than without biasing (black). If the bias further increases, the MFPT slows down round about a factor of two. Also, if the barrier height increases, the MFPT decelerates which becomes evident in Fig. 3.5 (b.1). The MFPT curves exhibit a dent around $z/R \approx 1$,

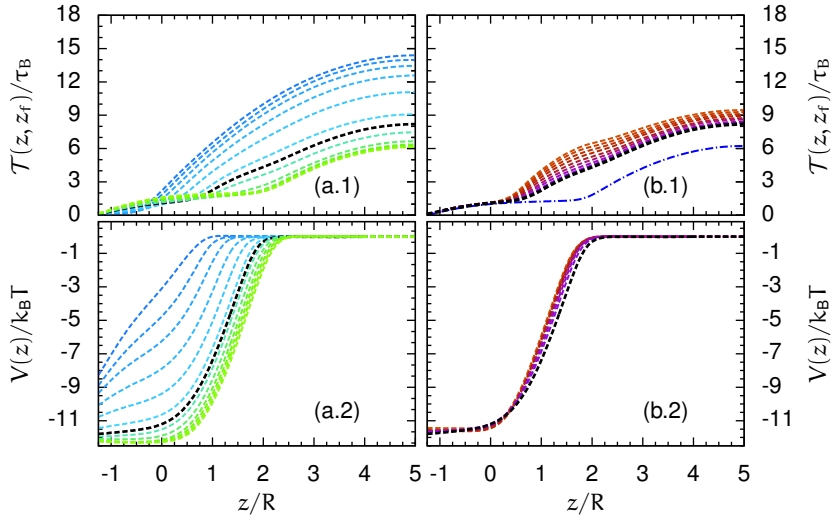

Fig. 3.5: The numerical results for the MFPT curves $\mathcal{T}(z,z_f)$ and PMFs $V(z)$ of ligand binding to the pocket depend on the changes in the water interface fluctuations if we apply the double-well variations from Fig. 3.3. Panel (a.1)/(a.2) shows the ligand's MFPTs/PMFs if the double-well biasing b changes. Panel (b.1)/(b.2) draws the ligand's MFPTs/PMFs if the double-well barrier height h changes. The blue dash-dotted line in panel (b.1) draws an example MFPT curve $\mathcal{T}_1^M(z,z_f)$ using Eq. (2.14) for the reference setting ($h = 1\ k_BT$, $b = 0$) with the respective PMF and constant friction. The color coding is adopted from Fig. 3.3.

which is enhanced with growing barrier height whereas for $z \lesssim 0$ all curves $\mathcal{T}(z,z_f)$ in panel (b.1) coincide.

Further, we analyze the PMFs of the ligand along z for all considered parameter values (h,b) using the WHAM [130, 131, 152]. For that purpose, we employ umbrella windows at restraining positions $z_1^r/R = -2$ to $z_{58}^r/R = 5.25$ in steps of $\Delta z/R = 0.125$. The spring constant is $k_z = 107.84\ k_BT/R^2$, which is the same but converted constant from MD simulations in Section 2.2.4. Fig. 3.5 (a.2) shows a strong dependence of the PMFs on the biasing parameter b. Besides small changes in shape, the attractive part of the PMFs essentially shifts towards smaller values of z if the interface bias shifts towards an increasingly wet pocket. In contrast, the double-well barrier height h negligibly tunes the ligand PMFs, which is revealed in Fig. 3.5 (b.2). These PMFs are equilibrium quantities and, thus, are essentially unaffected if mainly the evaporation rate of the pocket changes due to changes in h. This is especially noteworthy since the corresponding MFPT curves in panel (b.1) alter with h, suggesting that the effect on ligand binding times originates from modified water interface fluctuations.

In the case of a Markovian process, the PMF $V(z)$ together with a possibly spatially dependent friction $\xi(z)$ determine the n-th moment of the FPT distribution $\mathcal{T}_n^M$ as we depicted in Eq. (2.14) in Section 2.3. The blue, dash-dotted line in Fig. 3.5 (b.1) is such a theoretical Markovian estimate using constant ligand friction and the spatially

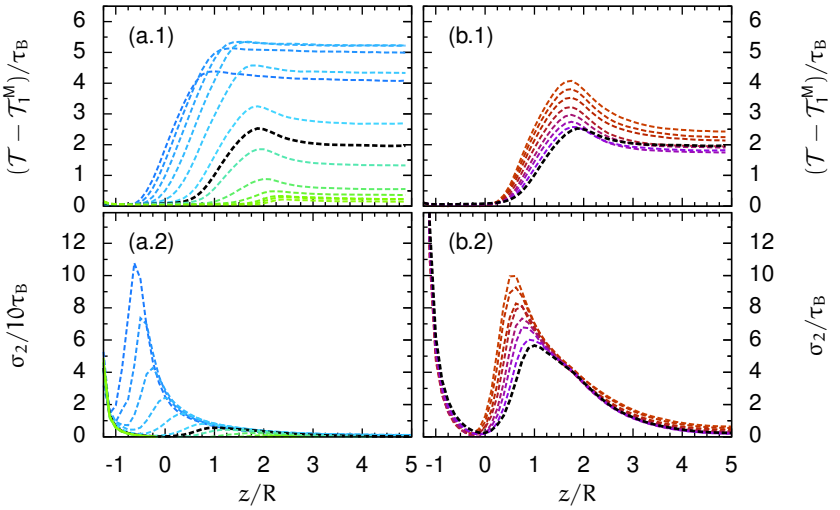

Fig. 3.6: The differences between numerical and theoretical results from Eq. (2.14) manifest that the Markovian assumption breaks down. Panels (a.1) and (a.2) show difference in the MFPTs and the memory index depending on the double-well bias b. Panels (b.1) and (b.2) draw both measures depending on the double-well barrier height h. Note that the effect is much bigger if the biasing parameter increases. Therefore the scale of (a.2) is $\sigma_2/10\tau_B$. The color coding is adopted from Fig. 3.3.

dependent $V(z)$ of the reference case $b = 0$ and $h = 1\ k_BT$. Hence, it should be compared to the black, dashed MFPT curve in the same plot. Both coincide only for negative z-values. Around $z \approx 1$, we observe effects, that can no longer be treated by Markovian kinetics nor by assuming constant friction.[1] There, a dent in the simulated MFPT curves qualitatively deviates to the theoretical prediction. For even bigger values of the reaction coordinate, the shapes of the MFPT curves of both methods only conform, but the Markovian solution represents overall faster association. In summary, the ligand kinetics can only be modeled by a pure Markovian description if the ligand is inside the pocket, which is by far least relevant to our investigations.

If we calculate the MFPT curves $\mathcal{T}_1^M$ in the Markovian picture from Eq. (2.14) for all considered cases of biasing and barrier height, we can directly compare the numerical and theoretical results. Accordingly, the difference $\mathcal{T}(z, z_f) - \mathcal{T}_1^M(z, z_f)$ is plotted in Fig. 3.6 (a.1) and (b.1). In all cases, the difference vanishes inside the pocket and maximizes just in front of the pocket mouth. It then plateaus to a constant positive value for large z-values. The difference $\mathcal{T} - \mathcal{T}_1^M$ in Fig. 3.6 (a.1) is very small if the bias is negative, namely if the pocket is rather dry. If the biasing parameter increases, and thus the water interface's distribution tends towards mainly wet pocket states, the difference accumulates to a maximum around $b = 2.8\ k_BT/R$, and

[1] We assume constant friction until the end of this section. Especially the calculations of $\mathcal{T}_1^M(z, z_f)$ and $\mathcal{T}_2^M(z, z_f)$ on the next page implement spatially independent friction.

alleviates again for even higher biasing. In Fig. 3.6 (b.1), deviations to the Markovian picture are enhanced by growing barrier height, hence, slower evaporation rates.

For further investigations of the break-down of Markovian kinetics and possibly accompanied memory, we also calculate a so-called memory index [153]

$$\sigma_2 = \frac{\mathcal{T}_2(z, z_f) - \mathcal{T}_2^M(z, z_f)}{\mathcal{T}(z, z_f)} \tag{3.7}$$

introduced by Hänggi *et al.* [153], who noted its possible value to ligand migration studies. $\mathcal{T}_2$ denotes the numerically sampled second moment of the FPT distribution and $\mathcal{T}_2^M$ refers to the respective measure from Eq. (2.14). Eq. (3.7) provides an additional spatially resolved measure of a random process' character. It indicates a process to be non-Markovian if the difference of the second moment of any random process to the second moment of the corresponding Markovian process does not vanish. In our cases in Fig. 3.6 (a.2) and (b.2), the memory index vanishes for ligand positions far from the pocket. Note, however, that σ_2 also takes non-zero values at intermediate positions, e.g., $z/R \sim 3$, where the ligand is still out of reach of the desolvation potential, i.e., $\zeta(z_s, z) > R$. Finally, the memory index peaks at the position $z/R \lesssim 1$. Subsequently, for $z/R < 1$, it steeply recedes to zero. Inside the pocket σ_2 diverges once more. Actual non-Markovian effects reoccur inside the pocket, which becomes evident from results and discussions in the following sections and in the Appendix B.2.

To this point, we corroborate that it is inept to assume constant ligand friction within a Markovian description of hydrophobic key-lock binding. The process rather indicates that non-Markovian contributions predominantly arise in the region where the ligand and the interface start interacting, which slow the ligand binding. Additionally, dissimilar MFPTs, but almost similar PMFs in Fig. 3.5 (b.1) and (b.2), respectively, suggest that water interface kinetics, i.e., evaporation rates, are sufficient to tune ligand binding kinetics, i.e., binding rates.

3.2.2 Spatially dependent friction

To illuminate the impact on ligand kinetics, we restrict the further analysis to systems with varying barrier height, but with unbiased pocket wetting, because of the previously found sheer kinetic impact in Fig. 3.5 (b.1) and (b.2). We replot the MFPTs for varying double-well barrier heights in Fig. 3.7 (b) together with the following results on spatially dependent ligand friction in panel (a).

To calculate the spatial dependence of the ligand friction $\xi(z)$, we employ umbrella setups, which are restraining the ligand at positions $z_1^r/R = 1$ to $z_{13}^r/R = 2.625$ in steps of $\delta z/R = 0.125$. The spring constant is the same that was used for PMF

calculations. (The choice of the interval was made by an initial coarse scan from restraining positions deep inside the pocket up to distances much further away.) We calculate the spatially resolved friction from the PACF in each umbrella window using Eq. (2.11) from Section 2.2.4. The solid curves in Fig. 3.7 (a) show Gaussian fits to the numerically sampled friction.[2] While the ligand is far away from the pocket, we observe our preset friction, but $\xi(z)$ peaks high in front of the pocket mouth. At this position, the ligand is subject to primary interactions with the water interface. Growing barrier heights and exponentially slowed double-well transition rates of the water interface, enhance the friction peak up to a factor of approximately 85. This indicates that the effect arises due to the ligand interacting with the bimodally fluctuating pseudo-particle. This can be pictured in the following way. While the water interface penetrates the pocket, the ligand, which is situated around $z \sim 2.0$, is only subject to the δ-correlated random force of our Brownian model. If the interface is in the outer well, in front of the pocket, the desolvation potential acts upon the ligand. Hence, the desolvation force acts as additional fluctuating force which introduces additional friction, considering the general statement of the FDT. The peaking friction occurs in the region in which fluctuations are most pronounced, where the water interface and the ligand might interact or not, and where the desolvation potential can essentially switch from *on* to *off* and *vice versa*.

Again, we calculate the MFPT curve $\mathcal{T}_1^{\mathrm{M}}(z, z_f)$ using the Markovian framework from Eq. (2.14), with the PMF $V(z)$ of the reference case ($h = 1\ k_B T$, $b = 0$), but now considering the spatially resolved friction $\xi(z)$ from PACFs. The result is plotted as the green dash-dotted line in Fig. 3.7 (b) and only coincides with the numerically sampled MFPT curve (black dashed) if the ligand is within the interval $z \in (-1.25, 0.5)$. Subsequently, a steep edge in the curve results in values, which exceed the numerically sampled results far outside, i.e., $z/R > 2$. So, on one hand, the Markovian prediction overestimates the numerical results if we use both spatially resolved profiles $V(z)$ and $\xi(z)$ from umbrella sampling; and, on the other hand, the numeric results are underestimated using only the PMF but constant friction. This well compares to the discrepancies that were observed in the results from MD simulations in Fig. 2.5 in Section 2.3 and Ref. [55].

For further comparison to the previous findings in MD simulations, we also calculate the spatially dependent profiles $\xi^{\mathrm{M}}(z)$, using Eq. (2.15). Note that $\xi^{\mathrm{M}}(z)$ uses the Markovian assumption, and is not the "proper" friction. However, this profile naturally reproduces the numerically sampled MFPT by construction. The dashed lines in Fig. 3.7 (a) show the curves for $\xi^{\mathrm{M}}(z)$, which also maximize in front

[2] The raw data and details on the fits are presented in Appendix B.2.

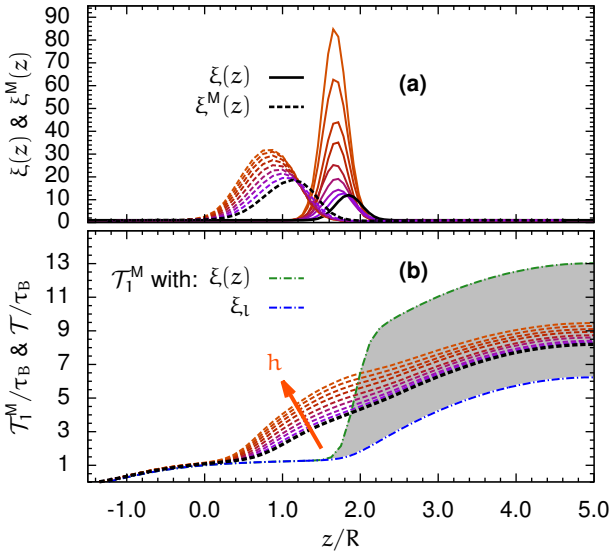

Fig. 3.7: Panel (a) plots the spatially resolved friction $\xi(z)$ from fits to the PACF data using Eq. (2.11) (solid) and from the MFPT data using Eq. (2.15) (dashed). Panel (b) draws again the MFPT curves from numerical calculations which were already shown in Fig. 3.5 (b.1). Here we also compare to the Markovian estimate $\mathcal{T}_1^{\mathrm{M}}(z, z_f)$ from Eq. (2.14) using the PMF and the spatially dependent friction from the PACFs, which is plotted as green dash-dotted line. The color coding is adopted from Fig. 3.3.

of the pocket mouth. In exact comparison to $\xi(z)$ from the PACF data, the results $\xi^{\mathrm{M}}(z)$ are spatially shifted towards the pocket and differ in peak height as well as in peak width.

Observing essential discrepancies between the local friction $\xi(z)$ from PACFs and the spatial kinetic profiles $\xi^{\mathrm{M}}(z)$ from the memoryless theory illustrates the location and strength of non-Markovian effects within our stochastic model. We emphasize that our model exhibits the similar non-Markovian effects as those resolved by explicit-water MD simulations in Refs [55, 101].

3.3 Generalized Langevin model

Having identified the fluctuations of the desolvation potential as the origin of local friction and non-Markovian kinetics in the ligand's reaction coordinate, we show next how the water fluctuations and ligand friction peaks relate. Further, this section shall formulate the proper stochastic characteristics when dealing with the ligand coordinate alone. For simplicity, we focus on the local conditions, namely while the ligand is close to the pocket where the friction maximizes during umbrella sampling. Nevertheless, this section elucidates a proper non-Markovian formulation to classify possible treatment of the problem with the conventional theory.

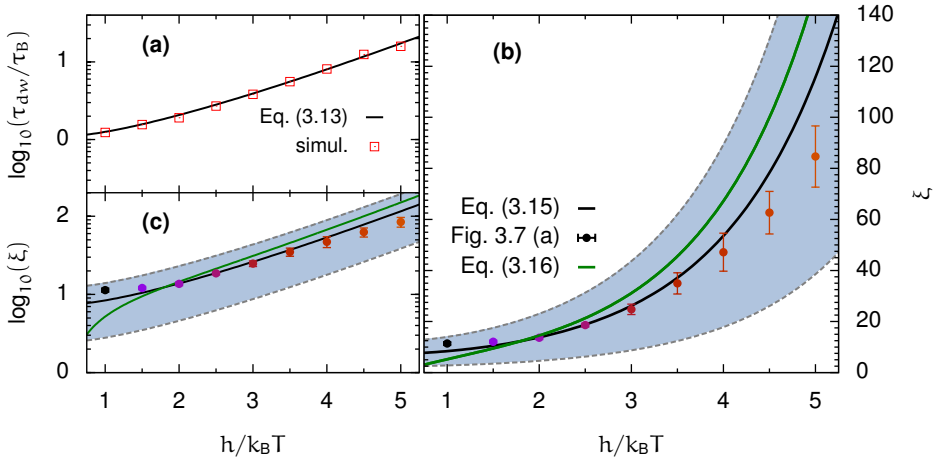

Fig. 3.8: Panel (a) plots the time scale $\tau_{dw}(h)$ of the interfacial motion in the double-well without ligand. The black line refers to Eq. (3.13) and the orange squares are the values obtained by exponential fits of the water interface's PACF from numerical calculations. Panel (b) shows $\xi_0(h)$ in Eq. (3.15) using $\epsilon = 0.36$ as the black line. The blueish shaded area, which is delimited by the gray dashed lines, indicates the range covered by Eq. (3.15) using the standard deviation $\Delta\epsilon = \pm 0.24$. The green lines show $\xi_0(h)$ from Eq. (3.16). The peaking values $\max(\xi(z))$ from Fig. 3.7 (a) are shown as circular symbols with the corresponding color coding which is adopted from Fig. 3.7. Panel (c) plots the same data as in (b) on a log-scale.

Generally, in the case of a known memory kernel $\xi(t)$ we can investigate the corresponding one-dimensional GLE

$$m\ddot{q}(t) = -\frac{\partial V_{eq}(q)}{\partial q} - \int^{t} dt' \xi(t-t')\dot{q}(t') + \mathcal{F}(t) \tag{3.8}$$

with mass m, equilibrium potential $V_{eq}(q)$, and a random force which fulfills the fluctuation-dissipation relation $\langle \mathcal{F}(t)\mathcal{F}(t')\rangle = 2\, k_B T m \xi(t-t')$. Simple systems of two coupled Langevin equations can be analytically contracted onto such an one-dimensional GLE and *vice versa*. A prominent example is that of an underdamped Brownian particle in a harmonic potential [111, 154, 155]. For the coupled system described by Eq. (3.6), analytic contraction from 2D to 1D is not feasible due to higher orders than harmonic coupling and the non-linearity from the double-well potential. Therefore, we reinterpret a method, which is usually used to expand a one-dimensional GLE to a set of two coupled equations without memory. We compare these equations to our system in Eq. (3.6) and approximate the friction from the local conditions of our key-lock model. We restrict the analysis to the location of the friction peaks discussed above and predict the peaking value $\max(\xi(z))$ as a function of the barrier height for the bimodally fluctuating force.

To this end we, reverse the approach from Pollak *et al.* [154, 156, 157]. This approach originally extends a one-dimensional GLE of reaction coordinate q, such

as Eq. (3.8), by an auxiliary variable x to receive two coupled equations. Each of the resulting equations then omits memory, and only the auxiliary variable is driven by a temporally δ-correlated random force $N(t)$. In particular, taking unit mass $m = 1$, the GLE in Eq. (3.8) is mapped onto the two dimensional, memoryless system

$$\ddot{q} + \frac{\partial V(q,x)}{\partial q} = 0 \tag{3.9a}$$

$$\xi_0 \dot{x} + \frac{\partial V(q,x)}{\partial x} = N(t)\,. \tag{3.9b}$$

The random force $N(t)$ is simply Gaussian white noise and is characterized by $\langle N(t)\rangle = 0$ and $\langle N(t)N(t')\rangle = 2k_B T\xi_0\ \delta(t-t')$. Two further requirements one for the memory $\eta(t)$ and another one for the coupling potential $V(q,x)$ must fulfill for the proper mapping:

(a) The kernel $\xi(t)$ must be represented by a sum of exponentials, and for this very example, we even assume

$$\xi(t) = \frac{\xi_0}{\tau} e^{-t/\tau} \equiv \Omega e^{-t/\tau}\,. \tag{3.10}$$

(b) The coupling between the auxiliary and the reaction coordinate should be harmonic such that

$$\frac{\partial V(q,x)}{\partial q} = \frac{dV_{eq}(q)}{dq} - \Omega[x - f(q)]\frac{df(q)}{dq}\,. \tag{3.11}$$

Now we wish to formulate these conditions for an approximate GLE model of the type of Eq. (3.9), hence Eq. (3.8), using the input to our original two-dimensional stochastic key-lock model from Eq. (3.6). Therefore, we focus on the position ϵ of the friction peak $\max(\xi(z)) \equiv \xi(\epsilon)$, which we observe in Fig. 3.7 (a). We expand the solvation force from Eq. (3.4) at fixed ligand-interface distance ϵ up to first order

$$F_{sol}(\zeta;\epsilon) = \pi\Gamma[\epsilon^2 - R^2] + 2\pi\Gamma\epsilon(\zeta - \epsilon) + \mathcal{O}(\zeta^2)\,, \tag{3.12}$$

which yields the harmonic contribution of our coupling potential between the water interface and the ligand. By comparison with Eq. (3.11), we identify the memory kernel constant $\Omega = \xi_0/\tau \equiv 2\pi\Gamma\epsilon$. As a striking difference we do not implement the double-well potential itself. We rather choose the memory time scale to be that of a

Brownian particle in a double-well. A compact approximate solution of that time scale is given by [158, 159]

$$\tau_{dw}(h) = \frac{\tau_B}{2\sqrt{2\beta h}} \frac{(e^{\beta h} - 1)}{2\beta h} (\pi\sqrt{\beta h} + 2^{1-\sqrt{\beta h}}) \,. \tag{3.13}$$

To confirm this approximation for the setups, which we previously considered, we probe the time scale of interface fluctuations in the double-well within numerical calculations without the ligand. Tuning the barrier height from $h = 1\ k_BT$ to $h = 5\ k_BT$ reveals that the approximate Eq. (3.13) is in very good agreement with our numerical results within the range of interest of h as it is plotted in Fig. 3.8 (a). Hence, the final GLE applies the harmonic potential coupling from Eq. (3.11) and the relaxation time from Eq. (3.13) such that

$$\ddot{q}_l(t) = -2\pi\Gamma\epsilon \int^t dt' \, e^{-(t'/\tau_{dw})} \dot{q}_l(t') + \mathcal{F}(t) \tag{3.14}$$

where $\mathcal{F}(t)$ fulfills the FDT. Comparison of its memory kernel with Eq. (3.10) determines the friction for the constructed system such that

$$\xi_0(h) = 2\pi\Gamma\epsilon \cdot \tau_{dw}(h) \equiv \Omega \cdot \tau \,. \tag{3.15}$$

Again, we focus on the position of the friction peak in Fig. 3.7. There, we want to approximate the average separation of the ligand to the water interface pseudo-particle. We numerically estimate the value of $\epsilon \approx 0.36 \pm 0.24$ while the ligand is constrained at the position of the friction peaks. The value ϵ is the mean of the distribution of distance ζ between the constrained ligand and the bimodally fluctuating interface where its error is the respective standard deviation. Details on these numerical estimates can be found in the Appendix B.3. Fig. 3.8 (b) and (c) compare the results of Eq. (3.15) with $\max(\xi(z))$ from Fig. 3.7 (a), where a strong resemblance of the non-Markovian model in Eq. (3.14) and the fully coupled key-lock binding model in Eq. (3.6) becomes evident. The circular symbols with error bars draw the maxima of the friction peaks $\max(\xi(z))$ from PACFs of the original key-lock model (see also Fig. 3.7). The black line draws the expression found for $\xi_0(h)$, i.e., Eq. (3.15). The blue shade indicates the error from variance calculations of ϵ.

In a more rigorous approach, we can also calculate the average interface-ligand separation ϵ utilizing the expression of the full energy landscape $V_{tot}(z_s, z_l)$ in Eq. (3.5). If we fix the ligand to $z_l/R = 1.7$, which is roughly the peaking friction position in Fig. 3.7, the Boltzmann factor $\exp[-\beta V_{tot}(z_s, 1.7)]$ serves as weighting

function to calculate the average distance $\epsilon = \langle \zeta(z_s, 1.7) \rangle = \langle 1.7 - z_s \rangle = 1.7 - \langle z_s \rangle$. In that case, the expression for the barrier dependent friction constant in the GLE system reads

$$\xi_0(h) = 2\pi\Gamma \left(1.7 - \frac{\int dz_s \; z_s e^{-\beta V_{tot}(z_s, 1.7, h)}}{\int dz_s \; e^{-\beta V_{tot}(z_s, 1.7, h)}} \right) \cdot \tau_{dw}(h) \tag{3.16}$$

where the dependence on barrier height of $V_{tot}(z_s, 1.7, h)$ is also emphasized. In Fig. 3.8, this expression for $\xi_0(h)$ is drawn with a green curve. This also suitably approximates the barrier height dependence of the peaking friction from the numeric results. The assumption of fixed $z_l/R = 1.7$, however, suits less the conditions of minor ligand position fluctuations within the numerical calculations. Hence, numerically extracting ϵ for Eq. (3.15) yields a slightly better approximation.

3.4 Concluding remarks

This chapter revealed the origin of increased friction and non-Markovian kinetics in hydrophobic key-lock binding, which were observed in previous work using explicit water MD simulations [55, 101] and other hydrophobic association processes [102, 103]. Here, we employed a minimalistic stochastic model that captured the fundamental features of hydrophobic key-lock binding: the pocket hydration bimodally fluctuated between dry and wet states due to capillary evaporation, and the ligand coupled to the fluctuating water interface by solvation free energy. Consequentially, the coupling was bi-directional as it did not only lead to strong ligand attraction but also implemented ligand-induced pocket drying.

First, we could prove that our model conforms with the findings of the previous explicit water MD simulations. The FPT analysis of ligand binding revealed non-Markovian contributions to the binding kinetics in front of the pocket, where the ligand association decelerated in our numerical calculations of the two-dimensional model. We could resolve the effective friction of the ligand in a reduced one-dimensionally mapped description by calculating the PACFs and found spatially enhanced friction in front of the pocket. The friction was peaking at positions where coupling to the water interface occurs most prominently. Also, qualitatively similar but quantitatively different results were found from the Markovian methodology in Eq. (2.15). These friction results spatially shifted in comparison to the PACF sampled friction, which was also previously observed in explicit-water MD simulations [55].

In our last step, we further corroborated the origin of friction by constructing Eq. (3.15) in a GLE model, which was restricted to the local conditions of the peaking friction in the original system of Eq. (3.6). The GLE utilized the coupling

desolvation potential between the ligand and the interface, as well as the time scale of the interface fluctuations as the relaxation time of the memory kernel. When we compared the friction of the constructed GLE to the locally peaking ligand friction from our originally two-dimensional stochastic model, both models very well coincided quantitatively. This agreement substantiated that non-Markovian kinetics within the two-dimensional stochastic system is the outcome of projecting the binding kinetics onto the ligand reaction coordinate alone.

4 Principles of tuning kinetics – the binding site

Given a clean tunability in the stochastic model of the previous chapter, we now return to atomistic MD simulations and address the question of how the pocket hydration and the coupled ligand kinetics vary in response to changes in hydrophobically confined binding pockets. The *in-silico* represented hydrophobic model applies the spherical ligand from Section 2.2.2 and the concave binding site in the non-polar wall from Section 2.2.3. The wall enables systematic modifications of the binding site topography and its water attraction strength. In particular, we tune the curvature and the depth of the binding pocket on an ångström length scale, which is likely to fall within the range of natural conformational fluctuations of single, flexible host molecules. We also scale the receptor hydrophobicity by modifying the water-wall dispersion interaction, thus, reflecting changes, which are likely to occur due to packing fluctuations in macromolecules, e.g., proteins. To obtain qualitative, generic trends, we then correlate the binding times with phenomenological hydrophobicity scales such as water fluctuations [133–140]. Finally, in a subsequent in depth analysis for a more quantitative theoretical assessment, we uncover further intricate details of the hydrophobic association process. In particular, we further corroborate our previously derived relation between solvent fluctuation time scales and locally enhanced ligand friction [160], and we apply a correction to reproduce binding times which derive from atomistic MD simulations.

4.1 Varying pocket physicochemistry

The principal hydrophobic pocket-ligand setup was presented in Fig 2.3 in Section 2.2.3, where we discussed the principle for our geometric modifications. Here, we make use of this adjustability, where the procedure is again exemplified by the virtual spherical probe volume in the upper left corner of Fig. 4.1: The virtual probe volume's inverse radius r_i determines the inner pocket curvature. The probe volume's perpendicular distance to the wall determines the pocket depth Δ. Hence, we define a given geometric setup by the tuple (r_i, Δ) in units of ångström. Our choices of pocket geometry are also tabulated in Fig. 4.1. The pocket depth varies in 1 Å steps between 4 Å and 10 Å, while efficacious variations of the pocket radius scan the values $r_i = 9$, 9.5, 10 and 11 Å. As reference geometry, we choose the (9.5, 7)Å setup.

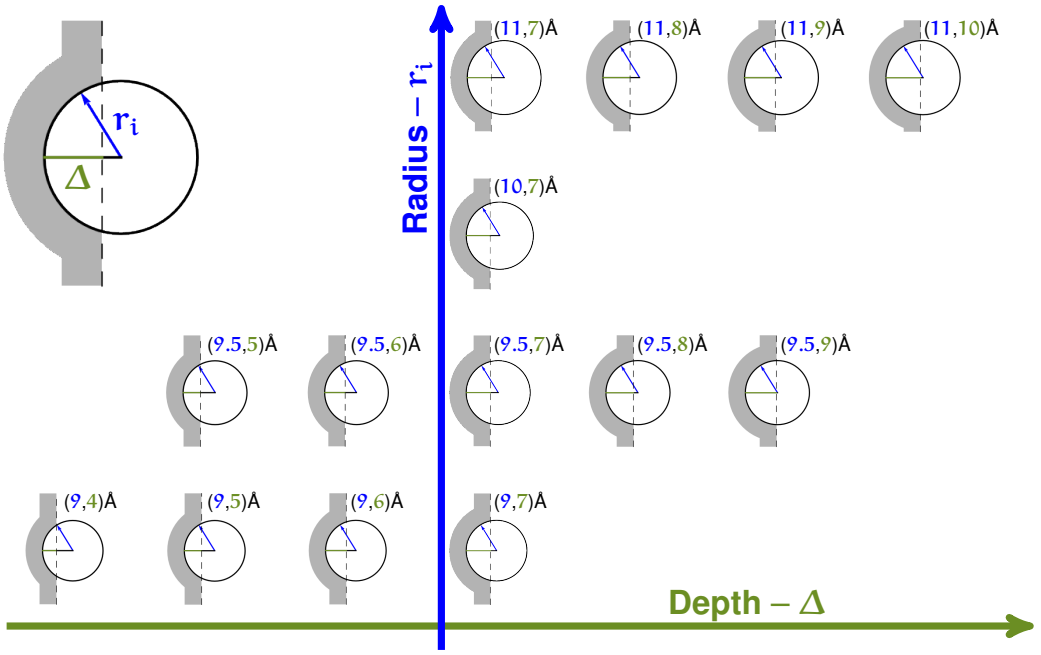

Fig. 4.1: The tabulated illustration shows the different geometry settings that we use. The gray pocket illustrations differ in curvature and depth as indicated by the tuples (r_i, Δ). We choose the geometry (9.5, 7)Å as reference system.

In an additional approach, we tune the hydrophobicity of the wall material by means of the wall-water interaction whereas the reference value is $\epsilon_{pw} = 39$ J mol^{-1}. For our reference geometry (9.5, 7)Å, we consider an additional parameter range from $\epsilon_{pw} = 29$ J mol^{-1} to 49 J mol^{-1} in steps of 5 J mol^{-1}. Thus, we change the wall-water Hamaker constant by approximately $\pm$13% and $\pm$25% [161]. Note that the interaction of the ligand neither to the water nor to the wall is modified by this procedure. This is also why we emphasize only the wall-water interaction parameter.

Let us only briefly summarize the principle simulation setups. The details of the two main simulation protocols for constrained and unconstrained simulations in the current chapter are described in Sections 2.2.4 and 2.2.5. All simulations are prepared as *NVT* ensembles at T = 298 K with a Nosé Hoover thermostat, that couple every picosecond in equilibration runs and every 10 ps in production runs. Note that only the solvent, i.e., the water, is coupled to the thermostat throughout the production runs. Hence, the kinetics of the ligand are unaffected by the thermostat. For technical reasons, we use constant volume simulations and the density of the water slab is individually tuned to (997.0 ± 0.3) g/dm^3 for each system by slight adjustments of the walls' separation. All simulations are performed with the GROMACS-4.6.3 simulation package [117].

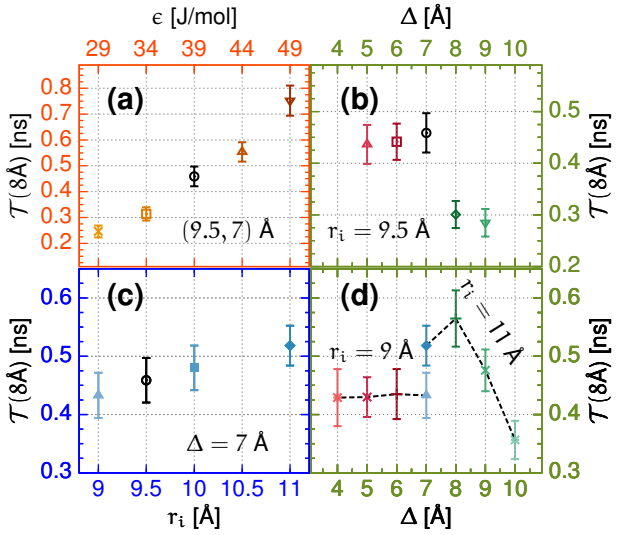

Fig. 4.2: (a) The mean binding time $\mathcal{T}$ from a distance of 8 Å increases if the water-wall dispersion interaction ϵ_{pw} is enhanced (i.e., lowered hydrophobicity) in the (9.5, 7)Å geometry. (b) If the depth of the pocket with $r_i = 9.5$ Å increases, the binding time abruptly accelerates, i.e. from $\Delta = 7$ Å to 8 Å. (c) Decreasing the pocket radius is slowing the binding kinetics while $\Delta = 7$ Å. (d) Changing the depth in the pockets with radii 11 Å and 9 Å also exhibits strongly accelerated binding times for deep pockets. The symbol and color coding of panels (a)-(d) is adopted throughout this chapter.

4.2 Binding time and pocket hydration

Since the process of the ligand approaching from far away to the pocket is standard diffusion, we focus in our following discussion on the mean binding time from $z = 8$ Å to $z_f = -2$ Å, i.e. $\mathcal{T}(8\text{Å}, -2\text{Å})$.[1] This starting position refers to a little more than one ligand diameter and is the location at which the kinetic coupling of pocket water occupancy fluctuations and ligand friction initiates.

The results in Fig. 4.2 show how the mean binding time $\mathcal{T}(8\text{Å})$ depends on the modulated pocket properties. In particular, the results show that an increasing water affinity to the binding site due to stronger wall-water dispersion attraction, as well as decreasing levels of confinement by increasing the pocket radius, result in a gradual slow down of the binding kinetics (Fig. 4.2 (a) and (c)). We achieve the opposite, but consistent, effect while increasing the pocket depth, thus, increasing the hydrophobic confinement (Fig. 4.2 (b) and (d)). Interestingly, we observe an abrupt reduction of the binding time once a certain critical pocket depth is reached. This sudden speed up is related to a rapid increase in the amplitude of the pocket-water occupancy fluctuations, as we see in the following paragraphs.

[1] In the following we abbreviate $\mathcal{T}(8\text{Å}) = \mathcal{T}(8\text{Å}, -2\text{Å})$. Remember that the pocket entrance defined the origin $z = 0$, which was shown in Fig. 2.3 of Section 2.2.3.

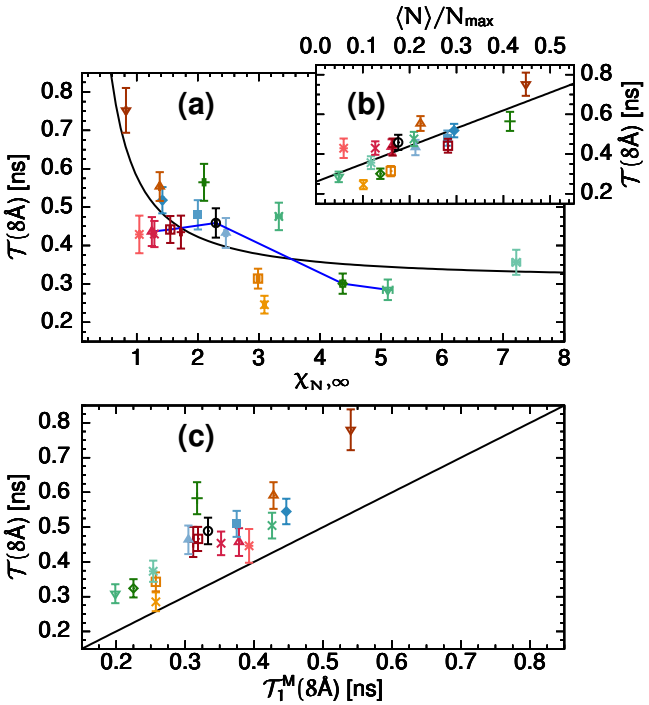

Fig. 4.3: The binding times $\mathcal{T}(8\text{Å})$ correlate with (a) the pocket water occupancy fluctuations $\chi_{N,\infty}$, characterizing the pocket's solvent repulsion, and (b) the mean water occupancy normalized by the maximal water occupancy, $\langle N\rangle/N_{max}$. Hence, low hydration and large water fluctuations give rise to faster binding of the ligand. The blue line connects the data from Fig. 4.2 (b) and illustrates that the abrupt binding speed up is associated with an abrupt increase of the solvent fluctuations. Solid black lines are empirical fitting laws of the form $A + Bx^{\alpha}$ with $A = 0.31$, $B = 0.27$ and $\alpha = -1.26$ in panel (a) and $A = 0.26$, $B = 0.90$ and $\alpha = 1$ in panel (b). In plot (c) the sampled mean binding time $\mathcal{T}(8\text{Å})$ is compared to the theoretical values $\mathcal{T}_1^M(8\text{Å})$ using Eq. (2.14) with constant friction $\xi_\infty = 0.4$ ns nm^{-2} and the PMFs. The horizontal distance to the black linear slope of one quantifies the discrepancy $\Delta\mathcal{T} = \mathcal{T}(8\text{Å}) - \mathcal{T}_1^M(8\text{Å})$. Symbol types and colors correspond to those of Fig. 4.2.

In a unifying view, the pocket's general hydrophobicity or degree of hydrophobic confinement results from the combination of chemical composition (water affinity/attraction) and topography (the level of confinement). Adopting a solvent-related perspective, we utilize the water fluctuations $\chi_{N,\infty}$ from Eq. (2.12) in Section 2.2.4 as an index for a pocket's degree of hydrophobicity.[2] As discussed in Section 2.2.4, this index relates to an isothermal compressibility [133] for large N, namely in the thermodynamic limit, because it relates to the free energy of desolvating the volume of interest [133–140]. Especially, if $\chi_{N,\infty}$ takes large values, the tendencies for pocket dewetting and drying increase, which is the case if hydrophobic confinement is enhanced. Similarly, one expects $\langle N\rangle$ to respond significantly to changes in the pocket physicochemical properties because of the sensitivities of dewetting or capillary evaporation to the geometry and water affinity of the nano-confinement [107, 133].

[2] $\chi_{N,\infty}$ absorbs its dependence on the ligand position into the subscript where $z = \infty$ means that the ligand is far away.

In Fig. 4.3 (a), we show the correlation of the mean binding times $\mathcal{T}(8\text{Å})$ from Fig. 4.2 with the fluctuation strength $\chi_{N,\infty}$. The correlation with average pocket occupancy $\langle N\rangle/N_{max}$ is shown in Fig. 4.3 (b). Here N_{max} is the individual maximal water occupation number. The comparison suggests a general trend that favorably wet, i.e., less hydrophobic, pockets increase the binding time on average and *vice versa*. The blue line illustrates the aforementioned abrupt speed up for deepened pocket geometries. Apparently, increasing the pocket depth slightly past a certain critical value can significantly enhance the water occupancy fluctuations. This manifests the transition to complete dewetting in an increasingly confining geometry. Intuitively, in such cases, water is more easily replaced by the ligand, in the sense of energetic costs and loses, which allows faster binding. In the extreme case of an all-time dry binding site $\langle N\rangle \to 0$ or a strongly dewetted binding site $\chi_N \to \infty$, the ligand can be expected to bind most quickly. On the contrary, if the pocket occupancy fluctuations decrease, pocket dehydration is less likely. Hence, for $\chi_N \to 0$ the binding time substantially increases, maybe even diverges, $\mathcal{T} \to \infty$, given the pocket is permanently occupied by water.

The above conclusions, indicate the link between the level of hydrophobicity and the *association speed*, which affirms the previously observed relation between hydrophobicity and binding *affinity* [56, 57, 87]. In order to investigate the directness of such a correspondence, we employ the equilibrium profiles of the potential energy landscape $V(z)$, plotted in Appendix C, with the Markovian approach to estimate the mean binding times $\mathcal{T}_1^M(8\text{Å})$ via Eq. (2.14). As a primary test to the above conclusions, we neglect the possibly spatial dependence of friction because it is hardly accessible in experiments. If we use Eq. (2.14) with a constant ligand resistance, i.e., $\beta\xi_\infty = 0.4$ ns nm^{-2}, we can estimate a lower boundary to the observed binding time [55, 160]. In Fig. 4.3 (c), we plot the observable $\mathcal{T}(8\text{Å})$ against the theoretical estimate $\mathcal{T}_1^M(8\text{Å})$. The comparison supports the general, qualitative conclusion that the binding time remains dominated by the binding affinity, whereas the latter is known to vary due to solvent repulsion effects, i.e., hydrophobicity [56, 87].

Importantly, however, we observe the systematic shift of the current theoretical predictions, $\mathcal{T}_1^M(8\text{Å})$, with respect to the simulated observables, $\mathcal{T}(8\text{Å})$, towards shorter binding times as indicated in Fig. 4.3 (c). This error, highly unsatisfying for quantitative rate predictions, arises from neglecting the friction's variations with the ligand-pocket separation. Unfortunately, the inclusion of spatially dependent friction into the analysis of hydrophobic association is also highly non-trivial [55, 102, 103, 160] as we repeatedly presented in all previous chapters. In the following section, we elaborate on the nature of the ligand friction and propose a model allowing

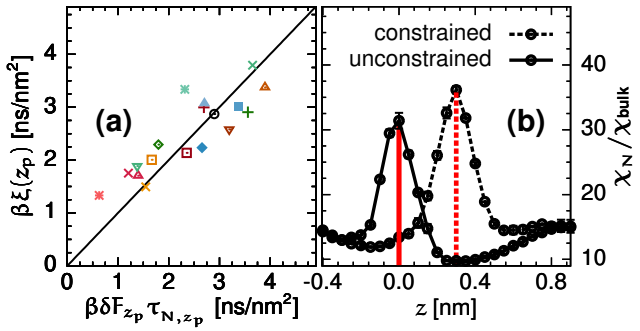

Fig. 4.4: (a) The product on the r.h.s. of Eq. (4.1), using simulation results of τ_{N,z_p} and δF_{z_p} at the peak position z_p, is plotted against values of $\beta\xi(z_p)$ from PACFs using Eq. (2.11). The black line draws a linear regression with a slope of one. Thus, the enhanced ligand friction is determined and tuned by changes in the pocket occupancy fluctuation timescale and the rate of change of the mean forces. The symbol and color coding is adopted from Fig. 4.2. (b) The occupancy fluctuations $\chi_N(z)$ are plotted in the example of the reference geometry (9.5, 7)Å. Its spatial dependence from unconstrained simulations (solid line) differs to that from constraining simulations (dashed line). The red solid and red dashed lines indicate the two different peaking position.

the inclusion of dissipative forces in a simple and quantitatively more satisfactory manner.

4.3 Kinetic coupling

By this section, the kinetic coupling of solvent fluctuations and ligand friction should be familiar since we depicted this topic in the all previous chapters. To prevent a repetitive discussion here, we summarize the results for the PMFs, friction profiles, pocket occupancy fluctuations, and their time scales in Appendix C. For The effective ligand friction $\xi(z)$ becomes spatially dependent and is locally enhanced prior to association. In particular, the friction was shown to maximize prior to ligand binding at a given position, which we call z_p (p for peak). We demonstrated in Chapter 3 that the strong solvent fluctuations depict the enhanced ligand resistance before association. While we focused on that region of locally maximized ligand friction at z_p, we derived the relation between the solvent fluctuation time scale $\tau_N(z_p)$ and the friction peak, $\xi(z_p)$.[3] Accordingly, we found that the magnitude of the ligand friction peak is linearly proportional to the water occupancy fluctuation time scale:

$$\xi(z_p) = \xi_{z_p} = \delta F_{z_p}\ \tau_{N,z_p} \tag{4.1}$$

with the coupling factor, δF_{z_p}, which was related to the harmonic potential coupling in Section 3.3. There, the second order perturbation of the desolvation potential determined the coupling constant of the water interface or the auxiliary variable in

[3] A positive correlation can also be observed in the respective profiles that are presented in Appendix C.

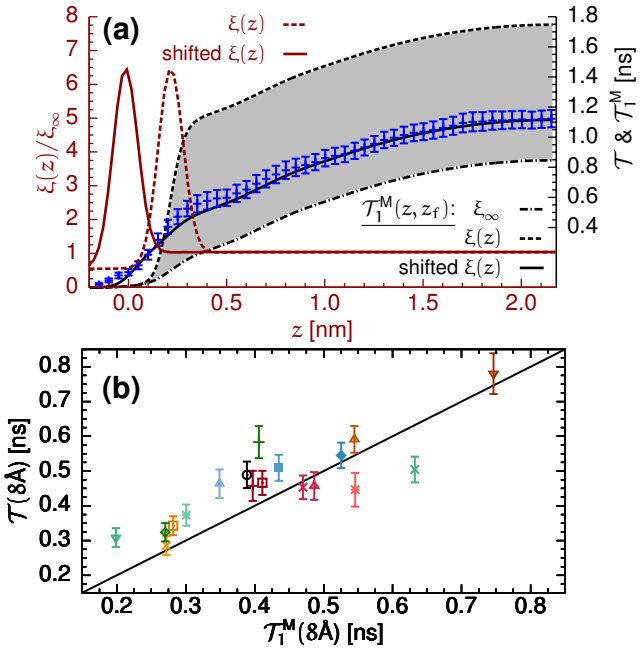

Fig. 4.5: Panel (a) illustrates how the ligand friction impacts the theoretical binding time estimation in the example of the most hydrophilic pocket, i.e., the (9.5, 7)Å-geometry with $\epsilon_{pw} = 49$ J/mol. The dash-dotted black line draws Eq. (2.14) with $\xi(z) = \xi_\infty$. The dashed black line plots the binding time using $\xi(z)$, shown as dashed reddish curve. The black solid line represents the binding time curve using the friction profile (reddish solid line) that accounts for the delayed solvent response by a negative spatial shift. The latter well coincides with the simulation sampled $\mathcal{T}(z, z_f)$, which is drawn as blue symbols with error bars. (b) The symbols plot $\mathcal{T}(8Å)$ from simulations against the respective estimate $\mathcal{T}_1^M(8Å)$ from Eq. (2.14) utilizing the accordingly shifted $\xi(z)$. The black line has a slope of one and quantifies the remaining discrepancies. The symbol and color coding is adopted from Fig. 4.2.

the GLE model of hydrophobic key-lock binding. Here, we say that the PMF serves as coupling potential, which is why we determine the coupling constant by the second spatial derivative of the PMF at the friction peak position, i.e., $\delta F_{z_p} = (\partial^2 V(z)/\partial z^2)_{z_p}$. The coupling force constant can be re-interpreted as the rate of change of the mean force. The relevance of Eq. (4.1) for our modified pockets, considered in the current chapter, is confirmed in Fig. 4.4 (a). The results of Eq. (4.1) correlate closely around the linear slope of one (black solid line) if we plot them against the simulation sampled friction peaks from PACFs via Eq. (2.11).

This rationalization of the kinetic coupling is observed in umbrella sampling simulations; however, the fluctuation peak is actually shifted by roughly 3 Å towards the pocket interior in Fig. 4.4 (b), if we compare water occupancy fluctuations in constrained and unconstrained simulations. Thus, the pocket's hydration response in constraining simulations barely materializes in the same way during the binding processes in unconstrained simulations. In other words, the solvent fluctuation and response manifest differently in the two simulation schemes. This also relates to

the non-Markovian effects shown by Setny [55] and our results in Chapter 3. If evaluating the friction from the MFPT and PMF utilizing Eq. (2.15), the friction peak shifts closer to the pocket. We already showed this in Figs 2.5 and 3.7 in Chapters 2 and 3, respectively.

Consequently, also in reference to the aforementioned kinetic coupling, we assume that the increased friction must shift to smaller values of z. Thus, as a straightforward solution to correct the error of the theoretical mean binding time estimates evidenced in Fig. 4.3 (c), we suggest to use the friction peaks from umbrella simulations, but to translate them by the individual shifts in $\chi_N(z)$ from constrained and unconstrained simulations. In Fig. 4.5 (a), we illustrate how the theoretical estimate for the MFPT in our most hydrophilic system with the (9.5, 7)Å geometry and a water-wall attraction of $\epsilon_{pw} = 49$ J mol^{-1} improves upon the correction of a spatially shifted friction peak. First, utilizing constant (bulk) friction $\xi(z) = \xi_\infty$ within Eq. (2.14), as before for Fig. 4.3 (c), yields a curve for $\mathcal{T}_1^{\mathrm{M}}(z)$ (black dash-dotted), which certainly underestimates the simulation sampled binding time curve (blue symbols with error bars). If we simply use the friction profile from umbrella simulations (reddish dashed line), the theoretical binding time (black dashed line) overestimates the simulation sampled binding times. Yet, shifting the friction (reddish solid line) and using it within Eq. (2.14) yields a result that coincides with the simulation sampled MFPT curve (blue symbols).

In summary, one can assess a lower and an upper boundary for the mean binding time (gray area in Fig. 4.5) whereas a more reliable computation incorporates the time transients that are not sampled in the ligand constraining simulations. Here, we have incorporated parts of the time transients by a straightforward shift of the friction profile. As one consequence, the impact of the friction peak is actually diminished. We discuss this impact in more detail in the next chapter.

Finally, we present how the mean binding time prediction improves for most systems, if Eq. (2.14) includes a position dependent friction that is shifted by the respective delay observed from occupancy fluctuations. In Fig. 4.5 (b), we show binding times measured from MD simulations, $\mathcal{T}(8\text{Å})$, versus the new theoretical predictions, $\mathcal{T}_1^{\mathrm{M}}(8\text{Å})$. The black line with a slope of one now corroborates that the binding time estimates improve if the friction peaks are shifted accordingly. Note, that possible peak height or width modulations due to solvent response upon binding are not accessible.

4.4 Concluding remarks

As a phenomenological observation in this chapter, we investigated how to steer the binding rate of hydrophobic key-lock/host-guest association while we systematically modified the physicochemical properties of our binding site and, thus, the resulting solvent-mediated effects. Our general and intuitively expected discovery was that the association speed positively correlated with an increasing degree of hydrophobicity [56, 57, 87]. However, our results less intuitively highlighted that host-guest interactions did not directly govern the binding kinetics, but rather indirectly, substantially mediated by solvent effects. Specifically, we pointed out that the degree of pocket wetting and the amplitude of water density fluctuations could serve as standalone descriptors of the binding times. Such a dependence on solvent behavior sometimes led to drastic effects, such as switching from slow to fast binding by increasing the pocket depth merely by 1 Å, which was enough to change the magnitude of solvent fluctuations abruptly.

In a second part, we dissected the maximized ligand resistance before binding due to the solvent fluctuations. We employed the fluctuation-dissipation relation, which we previously derived in Section 3.3 from our GLE model. This relation expressed the intimate coupling of the enhanced hydration fluctuation time scale of the pocket and the accompanying increase in ligand friction. It applied to all investigated physicochemical pocket variations.

Finally, we laid out that quantitative binding time predictions must account for the long time transients of the kinetic coupling. We found that the process of drying was delayed if we compared unconstrained and constrained simulations. Thus, probing the friction under ligand restraining conditions must subsequently incorporate a delaying spatial shift to advance the theoretical binding time estimates. In doing so, we essentially were able to improve the binding time predictions empirically.

5 Principles of tuning kinetics – the ligand side

In this chapter, we highlight how binding site hydration and ligand orientation influence the binding and unbinding times. Here, our first focus especially lies on the role of water in the reorientation pathway of ligands, which carry a non-spherical symmetry. We discuss the binding of aromatic compounds to pocketed and planar binding sites. We find that the solvation free energies of distinct sections of a ligand molecule steer the orientation pathway. In particular, the ligand orients so it can intersect the water interface while desolvating its most hydrophobic parts. In general, this behavior is pronounced on sharp water interfaces for instance at planar binding sites, whereas it is less pronounced on the dewetted water interfaces in the pocket. Moreover, we find that a given set of ligands favors two distinct orientation pathways. This bimodal behavior qualifies the ligand orientation as a fluctuating degree of freedom that adds to our kinetic coupling and local dissipative forces.

In the second part of this chapter, we discuss the kinetic coupling for the various ligands, and we return to the Markovian framework. By doing so, we conclude on our repeated discussion about the methodology that we can analyze our setup with existing Markovian methods. Particularly, we rescale the slow down into a free energy barrier such that it can be interpreted as an *kinetic barrier*. We depict how this kinetic barrier varies upon different ligands and find even more enhanced friction for the ligands with bimodally fluctuating orientation. We analytically discuss the impact of the additional kinetic barriers and find that the biggest implications arise for the unbinding process. A ligand's residence time in the bound state differs, for instance, by hundreds of microseconds and by more than a factor of five if we compare unbinding times which do or do not incorporate the kinetic barrier. Hence, we conclude that the solvent mediated kinetics and orientation pathways especially steer unbinding rates.

5.1 Varying ligands

In principle, we apply the same unconstrained and constrained simulation schemes from Sections 2.2.4 and 2.2.5. Here, we simply change to the various aromatic ligands from Fig. 2.2 in Section 2.2.2, namely ethylbenzene, toluene, benzene,

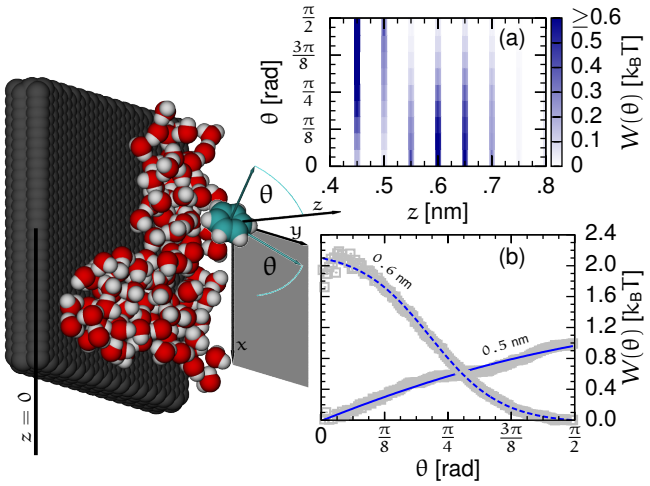

Fig. 5.1: The simulation snapshot illustrates benzene as ligand, the planar wall and part of the water. The binding process is constrained to one-dimensional diffusion, whereas the first crystal layer of the wall in contact with water defines $z = 0$. The ligand's orientation is quantified by the angle θ between the plane spanned by the ring atoms and the x-y plane (gray). (a) The dependence of the angular potential $W(\theta, z)$ on z illustrates the orientation pathway upon binding. The transition from a favorably perpendicular to a lateral orientation occurs on a narrow range from roughly $z = 7$ Å to 5 Å. Panel (b) exemplifies the sampled data for $W(\theta, z)$ as gray symbols including blue lined fits for $z = 0.6$ nm and $z = 0.5$ nm. Strikingly the perpendicular orientation is favored by over 2 $k_B T$.

phenol, and benzyl alcohol.[1] For all of these ligands, we run simulations with our reference (9.5, 7) Å geometry pocket from the previous chapter. Only for benzene, we additionally run simulations for the association to a planar wall.[2] Hence, we can compare the results of benzene binding into a pocket to the results of binding to a wall.

5.2 Reorientation

The simulation snapshots on the l.h.s. of Figs 5.1 and 5.2 illustrate how we define the benzene orientation by the angle θ. It is the angle between the benzene's plane spanned by its ring atoms and the x-y-plane (gray), which is the parallel plane to the wall.[3] Note that θ runs from 0 (lateral to the x-y plane) to $\pi/2$ (perpendicular to x-y plane) given the molecule's ring symmetry and, thus, its degenerate orientations. We sample the distribution of θ in each umbrella window and its Boltzmann inversion, the angular potential $W(\theta, z)$. Note that in this chapter $z = 0$ is differently defined than in previous chapters. For the wall, the first crystal layer, that is in contact with the water, defines the origin (see Fig. 5.1). For the pocket, the origin is defined

[1] All non-spherical ligands are also illustrated again in Fig. 5.3.
[2] The wall material is the same as for our molded binding sites.
[3] This angle is classically defined as angle between the normal vector of the benzene ring and the z-axis – the normal vector of the x-y plane.

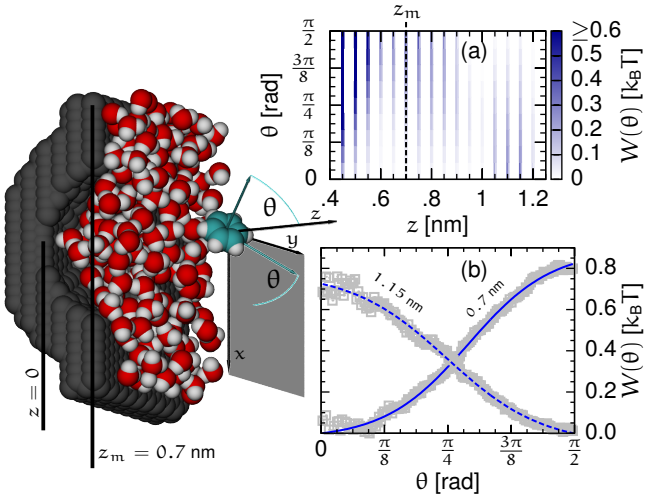

Fig. 5.2: The simulation snapshot illustrates benzene as ligand, a section of the hemispherical binding site and part of the water. The binding process is also constrained to one-dimensional diffusion along z. The pocket bottom is defined as origin $z = 0$ such that the pocket mouth is at $z_m = 7$ Å. Note, that this differs to the definition of the axis origin in our previous chapters. (a) The dependence of the angular potential $W(\theta, z)$ on z illustrates the orientation pathway upon binding. For 1.2 nm $> z >$ 1.0 nm a perpendicular orientation to the wall (and towards the water interface) is energetically favored whereas for $z <$ 1 nm it aligns with the x-y-plane. Panel (b) exemplifies the sampled data $W(\theta, z)$ as gray symbols including blue lined fits for $z =$ 1.15 nm and $z =$ 0.7 nm.

by the layer at the pocket bottom such that the pocket mouth is at $z_m = 7$ Å (see Fig. 5.2). This choice enables comparison of the results for these two binding sites.

Fig. 5.1 (a) and (b) draw $W(\theta, z)$ for benzene binding to the planar wall. Benzene undergoes an orientation pathway, which depends on the ligand separation to the wall. While the ligand is 6—7 Å away from the wall, the perpendicular orientation for $\theta = \pi/2$ is energetically favored by a little more than two k_BT. An example of $W(\theta, z)$ for $z =$ 0.6 nm is shown in Fig. 5.1 (b) where the simulation sampled data is represented as gray symbols and the dashed blue line is a fit. Thus, at these separations to the wall, benzene partly desolvates if it orients perpendicularly to the x-y plane. Proceeding to smaller z values, aligning parallel/lateral to the wall ($\theta = 0$) is favored by one k_BT and is sterically restrained upon final association. The example data (gray symbols) and fit (blue solid line) of the angular potential for $z =$ 0.5 nm are again shown in Fig. 5.1 (b).

In Fig. 5.2 (a) and (b), we see that the benzene association to the pocketed binding site is qualitatively similar but the reorientation from perpendicular to lateral occurs on a much wider range in z. The perpendicular orientation is energetically favored while the ligand is 1—1.2 nm away from the pocket bottom. Also, in comparison to the observations for the planar binding site, the perpendicular orientation is only slightly favored by little more than half a k_BT at $z =$ 1.15 nm.

This is exemplified in Fig. 5.2 (b). For even closer ligand-pocket separations around $z = 0.7$ nm the lateral orientation is favored by more than half a thermal energy as shown in panel (b), such that the benzene is already aligned with the pocket bottom before it is sterically restrained.

On one hand, binding of benzene to a hydrophobic binding site involves an energetically favorable, but possibly disadvantageous, perpendicular orientation. On the other hand, this pathway can be advantageous for slit like binding sites for which a ligand must enter and bind perpendicularly. The overall orientation pathway seems to be steered by desolvation of the ligand: the molecule can partly desolvate while orienting its edge towards the water interface, i.e, taking a perpendicular angle to the x-y plane. In the case of a hemispherically molded binding pocket, the strongly dewetted water interface allows even earlier desolvation and reorientation upon binding. As a consequence, the disadvantageous pathway is "smeared" at dewetted surfaces regarding the extended spatial range and lower energy barriers. Overall, this hydrophobically steered reorientation process particularly depends on the level of dewetting of a binding site. Stronger dewetting can assumingly loosen the rigid reorientation pathway. Also, different ligands exhibit different reorientation pathways where especially hydrophilic patches intuitively remain solvated as long as possible.

Our additional observations on reorientation pathways cover the association of our remaining aromatic compounds benzyl alcohol, phenol, toluene, and ethylbenzene to the pocketed binding site only (see bottom of Fig. 5.3). All of these ligands comprise an aromatic ring onto which an additional residue is attached. This residue breaks the ring symmetry, which was assumed for the benzene ring. In contrast to the analysis of benzene orientation, we define a new angle ϕ. This angle between the respective residue and the z-axis is schematically illustrated in the upper sketch of Fig. 5.3. Note that the unique mapping onto ϕ ranges from 0 to π, which is the necessary descriptor range to distinguish all orientations. If the angle is zero, the residue points into the water, away from the pocket. If the angle is equal to π, the residue points into the pocket, away from the water. For the benzene ring these two example angles were degenerate. Now, we sample the potential $U(\phi, z)$ from the orientation distributions in all umbrella windows in the same way we obtained the potential $W(\theta, z)$.

In Fig. 5.3 (a), the curves show examples for $U(\phi, z)$ from ethylbenzene and toluene constrained at $z = 11$ Å. Both examples exhibit a barrier around $\phi = 3\pi/8$, whereas the barrier is smaller in the case of toluene. This angular potential seems to be characteristic for these two aromatic ligands, which comprise an alkyl residue. If

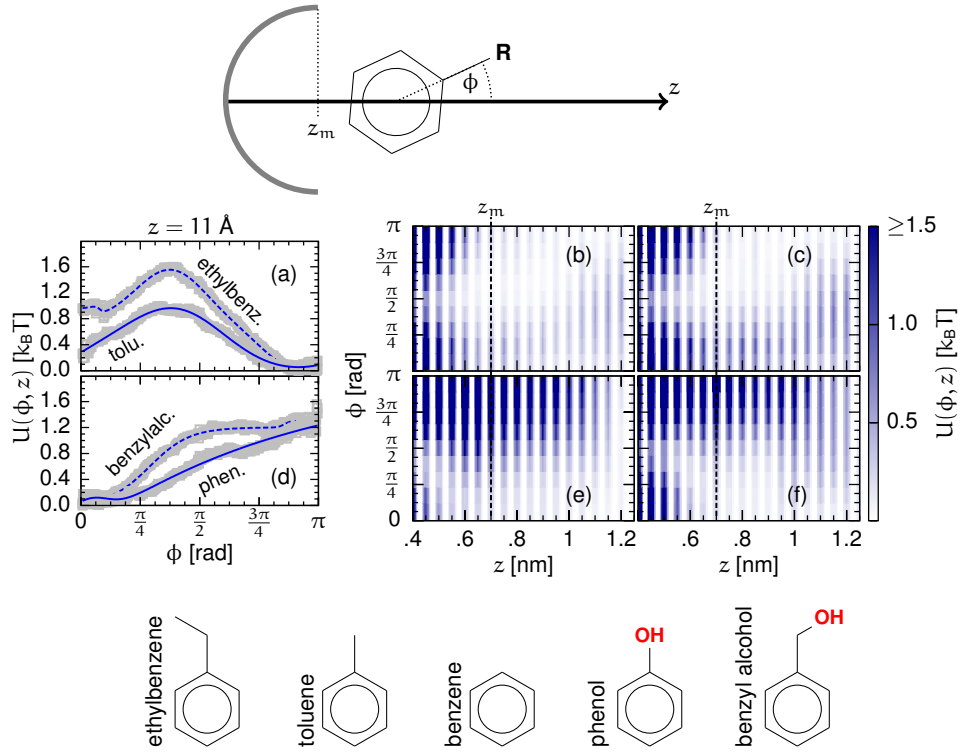

Fig. 5.3: The upper sketch schematically represents the pocket and ligand connected by the z-axis, whereas the pocket mouth at $z_m = 7$ Å is labeled. The angle ϕ is taken to be the angle between the respective ligand's residue R and the z-axis. Plot (a) shows two examples for the angular potentials $U(\phi, z = 11\text{ Å})$ from umbrella sampling of ethylbenzene (dashed line) and toluene (solid line). The gray symbols plot the sampled data and the blue lines represent smooth interpolation functions. On the r.h.s., the color map plots show the angular potentials from all umbrella windows for (b) toluene and (c) ethylbenzene. For these ligands, we observe a bimodal orientation pathway in the range $8\text{ Å} < z < 12$ Å. Panel (d) also shows the examples of $U(\phi, z = 11\text{ Å})$ for benzyl alcohol (dashed line) and phenol (solid line). Also for (e) phenol and (f) benzyl alcohol we show the full angular potentials $U(\phi, z)$ as color map plots. These two ligands preferably solvate there hydrogen bonding residue such that $\phi = 0$ is energetically favored. The lower sketches again illustrate the non-spherical ligands, which were originally shown in Fig. 2.2 in Section 2.2.2

toluene and ethylbenzene are at intermediate positions, they partly solvate either the aromatic ring (minimum at $\phi = \pi$) or the alkyl group (minimum at $\phi = 0$). Both of these orientations yield an energetic gain over an unfavored tilted orientation around $\phi = 3\pi/8$. Nevertheless $\phi = \pi$ is globally favored because the solvation of the aromatic ring yields higher energetic contributions from the electrostatic energy due to its partial charges. Fig. 5.3 (b) and (c) show that the bimodal orientations of toluene and ethylbenzene range from $z = 1.2$ nm to 0.9 nm. Finally, in the bound state around $z = 0.4$ nm, ethylbenzene and toluene are sterically hindered to take orientations other than $\phi \approx \pi/2$.

The orientation pathway is again very different for phenol and benzyl alcohol. In Fig. 5.3 (d), the potentials $U(\phi, z = 11\text{Å})$ for benzyl alcohol and phenol exhibit a

clearly favored orientation for $\phi = 0$. Both of these ligands have hydrophilic residues, which are terminated with hydroxyl groups. These impose two potential hydrogen bonding sites, namely the oxygen and the associated hydrogen atom. Hence, the favorable solvation of the hydroxyl group yields a strong orientation to $\phi = 0$. This is enhanced at even closer distances z, such that the barrier around $\phi = \pi$ increases to several thermal energies. In Fig. 5.3 (e) and (f) for phenol and benzyl alcohol, we can see that the energy to take an almost perpendicular orientation exceeds the plotted scale of 1.5 k_BT. Overall, these plots make evident that benzyl alcohol and phenol favorably solvate their hydroxyl groups, while they approach the pocket. Finally, both ligands orient towards $\phi = \pi/2$ if they are sterically forced to do so in the bound state.

We conclude that the orientation of all ligands is driven by solvation free energy such that the parts, which have the highest energetic costs upon hydration, are primarily desolvated. In contrast to previous density functional theory (DFT) studies on depletion and a resulting entropic torque for aspherical particles, we demonstrate that the principle reorientation pathways are individually steered for specifically considered chemical compounds. Since the pathways are driven by solvation free energy, a comparably termed *hydration* or *hydrophobic* torque generally comprises enthalpic as well as entropic contribution. After all, the phenomenon of a hydration torque is as diverse as the possibilities in constructing chemical compounds; however, it can be intuitively steered by knowing the chemical subunits of a ligand and the subunits' relative solvophobicity.

5.3 Kinetic barrier and rescaling

In Fig. 5.4, we plot the MFPT curves and the PMFs for all aromatic compounds as well as for the spherical ligand from previous Chapters 2 and 4. The upper panels (a.1) and (a.2) show the curves $\mathcal{T}(z, z_f)$, which make evident that ethylbenzene binds slowest. Though, this is a natural result because the bulk friction constant for ethylbenzene is the largest (see Appendix D.1). If we scale the MFPT curves by the respective bulk friction constants in panel (b.2), the curves for benzene, phenol, and benzyl alcohol coincide. The scaled binding times of ethylbenzene and toluene are slightly decelerated in comparison to the benzene binding times. This deceleration is counterintuitive when looking at panel (c.2). There the PMFs of ethylbenzene and toluene suggest stronger attraction and binding for these two ligands. Moreover, we compare benzene binding to the planar wall and pocket in panels (a.1) and (b.1). It binds comparably slower to the wall than to the pocket, which is in qualitative agreement with the PMFs in panel (c.1). Nevertheless, the PMF for the spherical

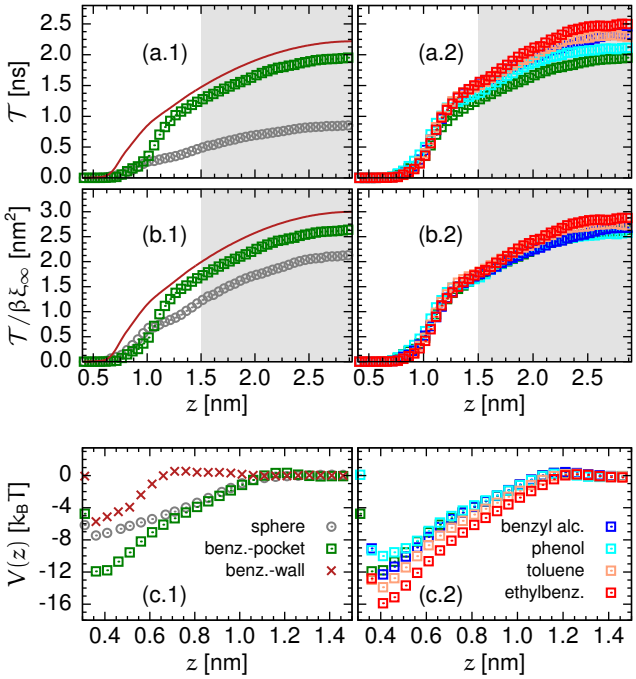

Fig. 5.4: The various MFPT curves in panels (a.1) and (a.2) naturally deviate because each ligand has a different friction coefficient. The red line is actually a theoretical prediction using Eq. (2.11) for the benzene association to the wall. In Appendix D.2 we depict that this theoretical prediction is sufficient. If scaling the MFPT curves by the respective bulk friction values $\beta\xi_\infty$ in panels (b.1) and (b.2) the effective differences become more obvious. Clearly benzene (green squares) is binding slower to the pocket than the spherical ligand (gray circles). Benzene binding to the wall (red line) is even slower. The friction normalized MFPT curves for the remaining aromatic compounds coincide with that of benzene. Only slight effective decelerations are observable for toluene (orange squares) and ethylbenzene (red squares). Panels (c.1) and (c.2) draw all PMFs in the interesting spatial range only. This spatial range corresponds to the white area in the (a.x) and (b.x) panels.

ligand suggests effectively slower binding to the pocket than benzene; however, this is not the case in the scaled MFPT curves in panel (b.1).

Throughout this thesis, we discussed that the pocket water density fluctuations yield additional friction forces that slow the binding. Particularly, we showed in Chapter 4 that important time transients could not be captured in ligand constraining simulations, which essentially impose an (artificial) equilibrium situation. However, we could show how the friction can be shifted accounting for a shift in the water fluctuations if we compared the results from constrained and unconstrained simulations. In principle, the direct coupling of fluctuations and friction, which we derived in Chapter 3, suggested that the friction should shift if the fluctuations shift. Hence, the shifted friction then improved the MFPT predictions in Chapter 4. The shifted friction peak position would similarly coincide with the peak positions of the profiles that could be derived from the Markovian method in Eq. (2.15). Hence,

we return to these profiles. Given the PMFs and MFPT curves, we can evaluate the "friction" profiles $\xi^{M}(z)$ from Eq. (2.15). In our case, long time transients of the non-Markovian effects are also incorporated in these profiles. In principle, we call it the *kinetic* profile $\xi^{M}(z)$, which naturally reproduces the MFPT.

Fig. 5.5 (a.1) shows the normalized kinetic profile $\xi^{M}(z)/\xi_{\infty}$ for the three previously discussed association processes. The results for benzene binding to the wall confirm that $\xi^{M}(z)$ can be assumed to be constant in that case. We assume that, at least for benzene, the rotational degrees of freedom are not coupling to the binding kinetics. In principle, the additional degree of freedom – orientation – could give additional features in the kinetic profiles. Yet, we assume that, also upon binding to the pocketed binding site, the kinetics are uninfluenced in the case of benzene. Fig. 5.5 (a.1) also illustrates the common and dominant peaking feature of $\xi^{M}(z)$, which we model and fit by

$$\xi(z) = \xi_{\infty} + \Delta\xi e^{-(z-z_p)^2/\sigma^2} \tag{5.1}$$

whereas the peak height $\Delta\xi$, position z_p, and width σ define a Gaussian shaped peak that adds to the bulk friction constant ξ_{∞}. In comparison to the spherical ligand, the kinetic profile of benzene peaks wider and shifts further into the bulk by roughly one ångström, which makes its tail reach to roughly $z \approx 1.2$ nm. This position coincides with the position where the perpendicular orientation is favored (see Fig. 5.2 (a)). Hence, as soon as the extended side of benzene, i.e., the edge, can reach towards the fluctuating interface, it is exposed to the increasing friction. In the following section, we see that this minor difference of the peak positions gives slower benzene binding; however, the PMF of the spherical ligand to the same pocket is less strongly binding.

The kinetic profiles for the remaining aromatic compounds ethylbenzene, toluene, phenol, and benzyl alcohol are shown in Fig. 5.5 (a.2) where they are compared to the replotted profile of benzene. The compounds phenol and benzyl alcohol are extended by an hydroxyl group and a methanol group, respectively, thus offering hydrophilic patches. Their size is elongated compared to the benzene ring but their kinetic profiles compare to the one of benzene. We conclude that, since the orientation pathways of these two ligands dominantly expose their aromatic ring to the pocket, the hydration fluctuations yield a similar kinetic profile as that of benzene.

Ethylbenzene and toluene are hydrophobic compounds, which are made up of a conjugated carbon ring that is extended by an ethyl and a methyl group, respectively. Their kinetic profiles in Fig. 5.5 (a.2) are more enhanced and reach even further

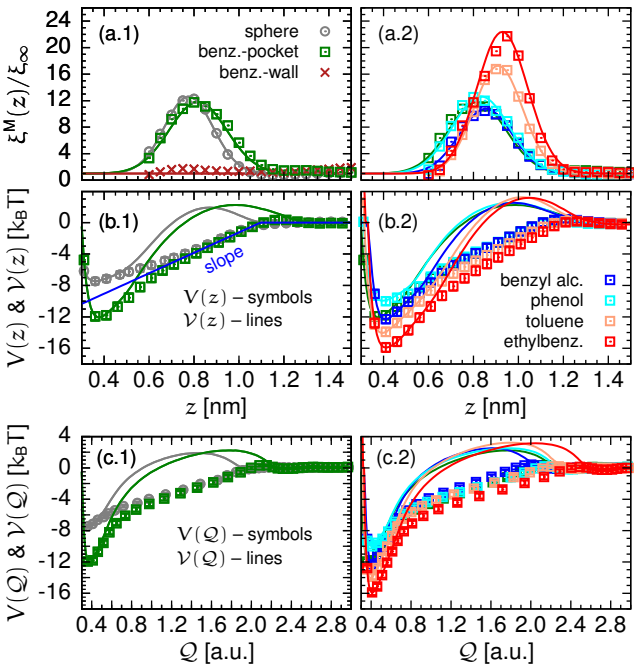

Fig. 5.5: The kinetic profiles in panels (a.1) and (a.2) maximize for all aromatic compounds binding to the pocket. The accordingly colored lines represent Gaussian fits of Eq. (5.1). If benzene, however, associates to the wall the friction can be well assumed to be constant. In (b.1) and (b.2) the original PMFs $V(z)$ are replotted as symbols. The rescaled potentials $\mathcal{V}(z)$ from Eq. (5.2) are shown as accordingly colored lines. The blue line demonstrates that the primary attractive slopes of both PMFs are similar. They exhibit the additional kinetic barriers due to Eq. (5.2). Panels (c.1) and (c.2) show $V(\mathcal{Q})$ (symbols) and $\mathcal{V}(\mathcal{Q})$ (lines) along the rescaled coordinate.

into the bulk. For those ligands, we observed an additional degree of freedom – orientation – that bimodally fluctuates. We conclude that the bimodally fluctuating orientation adds to the bimodally fluctuating pocket hydration and, thus, to the peak of the kinetic profile. In essence, the two degrees of freedom fluctuate bimodally, which yield additional dissipative forces in the one-dimensional description.

In order to judge the qualitative impact of the kinetic profiles, we rescale to a new reaction coordinate $\mathcal{Q}(z)$, as suggested by Hinczewski *et al.* [146]. The coordinate can be arbitrarily rescaled as long as the energy landscape, in our case the PMF, is consistently rescaled. We choose a rescaling such that the friction is scaled to a spatially constant value of one. The rescaled coordinate is determined by $\mathcal{Q}' = \mathrm{d}\mathcal{Q}/\mathrm{d}z = \sqrt{\xi^{\mathrm{M}}(z)/\xi_\infty}$ and the PMF must be consistently rescaled such that

$$\begin{aligned}\mathcal{V}(\mathcal{Q}(z)) &= V(z) + (\beta)^{-1}\ln(\mathcal{Q}'(z)) \\ &= V(z) + (2\beta)^{-1}\ln(\xi^{\mathrm{M}}(z)/\xi_\infty)\,.\end{aligned} \tag{5.2}$$

In Fig. 5.5 (b.1) and (b.2), we plot the rescaled potentials against the original coordinate z. In panels (c.1) and (c.2) the new coordinate $\mathcal{Q}$ is used. The rescaled potentials exhibit additional *kinetic barriers*, which naturally arise from the peaks in the kinetic profiles. Additionally, the rescaled coordinate is effectively stretched, thus, incorporating part of the effective slow down. In comparison to the barrier for the spherical ligand (gray line), the kinetic barrier for the aromatic compounds (colored lines) shift further away from the pocket, namely to increasing values of $\mathcal{Q}$. In general, the further the barrier shifts down the attracting slope the smaller its impact because the attracting slope diminishes the repulsive slope on the r.h.s. of the barrier. In particular, the repulsive slopes of the kinetic barriers of the aromatic compounds reach across the inset of attraction, which makes the barriers more effective. This argument is consistent with the MFPT data in Fig. 5.4 (b.1) and (b.2), where benzene and other aromatic compounds bind slightly slower than the spherical ligand. The binding times of ethylbenzene and toluene are even slower. The binding speed of phenol and benzyl alcohol, however, are similar to the binding times of benzene, which is consistent with the overlapping kinetic barriers in Fig. 5.5.

In summary, the size and nature of a ligand can shift and tune the height of the peaks in the kinetic profiles and the resulting kinetic barriers. In general, the number of bimodally fluctuating degrees of freedom increase the kinetic barriers, which slow the binding times. Here, we only discussed this phenomenologically with a scientist's intuition for the shapes of energy landscapes. In the following section, we approach our phenomenological arguments in an analytical discussion where we explore the full range for the possible impact of the kinetic barriers.

5.3.1 Impact of kinetic barrier

The following discussion is reduced to the fundamental features of the dissipative forces, i.e., the kinetic profile, and the PMF. We simplify $\xi^{M}(z)$ to its dominant feature, the fitted Gaussian peak from Eq. (5.2). The common and dominant feature in the PMF is its significantly attracting slope with roughly $f = 13\ k_BT/nm$ (blue line in Fig. 5.5 (b.1)). We simply take $V(z) = f(z - \bar{z})$ for $z \leq \bar{z}$ and $V(z) = 0$ otherwise. The position $\bar{z}$ denotes the inset of a constant attraction with strength f. The simplified potential and the kinetic profiles are illustrated in Fig. 5.6 (a). The gray and the green lines refer to the fitted peaks in Fig. 5.5 (a.1). The contribution of the friction peak, namely the contributions of the second summand on the r.h.s. of Eq. (5.1), to the binding time is given by

$$\Delta\mathcal{T} = \beta\Delta\xi \int_{z_f}^{z} dz' e^{\beta V(z') - \frac{(z'-z_p)^2}{\sigma^2}} \int_{z'}^{z_{max}} dz'' e^{-\beta V(z'')} \tag{5.3}$$

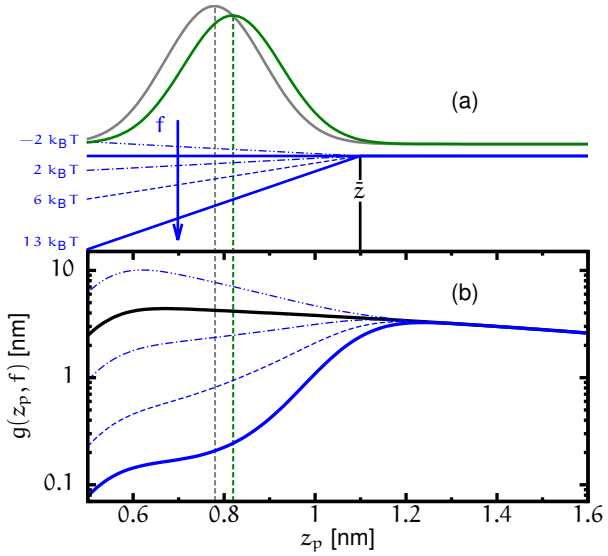

Fig. 5.6: (a) The dominant features of the ligand binding process are the potential $V(z) = f(z-\bar{z})$ (blue), that strongly attracts the ligand given $z \leq \bar{z}$, and a Gaussian friction peak (gray and green) modeled by Eq. (5.1). Various other potential slopes are sketched as blue lines with various line types. (b) The factor $g(z_p, f)$ in Eq. (5.6) and Eq. (5.7) strongly depends on the friction peak position. While the negative shift $z_p - \bar{z}$ decreases, $g(z_p, f)$ and thus the impact of the friction peak decreases.

whereas the stepwise definition of $V(z)$ has yet to be evaluated. Note that a discretization of the potential by a Heaviside Step function cannot be incorporated without contradiction but rather requires the scientist's intuition, which we describe here. Piecewise evaluation of the inner integral $\mathcal{I}(z)$ comprises a trivial case while $z > \bar{z}$, i.e., an integral over unity, and an integral over the Boltzmann factor $e^{-\beta V(z'')}$ while $z \leq \bar{z}$. The result remains piecewise defined, such that

$$\mathcal{I}(z) = \begin{cases} \frac{1}{\beta f}\left[e^{-\beta f(z'-\bar{z})} - 1\right] + (z_{\max} - \bar{z}) & \text{for } z' \leq \bar{z} \\ z_{\max} - z' & \text{for } z' > \bar{z}\,. \end{cases} \tag{5.4}$$

Note that the piecewise definition $\mathcal{I}(z)$ could be lost if $V(z)$ would have been discretized by a Heaviside Step function. Now, the outer integral reads

$$\Delta\mathcal{T} \propto \int_{z_f}^{z} dz' e^{-(z'-z_p)^2/\sigma^2} e^{\beta V(z')} \times \mathcal{I}(z') \tag{5.5a}$$

$$= \int_{z_f}^{\bar{z}} dz' e^{-(z'-z_p)^2/\sigma^2} \left[\frac{1 - e^{\beta f(z'-\bar{z})}}{\beta f} + (z_{\max} - \bar{z}) e^{\beta f(z'-\bar{z})}\right] \tag{5.5b}$$

$$+ \int_{\bar{z}}^{z} dz' e^{-(z'-z_p)^2/\sigma^2} \times (z_{\max} - z') \tag{5.5c}$$

where we used the piecewise definitions of $V(z)$ and $\mathcal{I}(z)$ to split the integral from z_f to z into an integral from z_f to $\bar{z}$ and another one from $\bar{z}$ to z. In particular, the inverse Boltzmann factor $e^{\beta V(z')}$ in Eq. (5.5a) is pulled into the square brackets in

Eq. (5.5b) (and it is one in Eq. (5.5c)). Completing the squares, if necessary, all integrals can be related to a Gaussian/Euler-Poisson integral. If we neglect second order terms of the peak width $\mathcal{O}(\sigma^2)$, the full solution reads

$$\begin{aligned}
\Delta\mathcal{T} \;\propto\; & \left\{ \frac{1}{\beta f}\left(\operatorname{erf}\left[\frac{z_p - z_f}{\sigma}\right] - \operatorname{erf}\left[\frac{z_p - \bar{z}}{\sigma}\right]\right) && \text{(5.6a)}\\
+ \; & \left(-\frac{1}{\beta f} + (z_{max} - \bar{z})\right) e^{\beta f (z_p - \bar{z})} \operatorname{erf}\left[\frac{\beta f \sigma}{2} + \frac{z_p - z_f}{\sigma}\right] && \text{(5.6b)}\\
- \; & \left(-\frac{1}{\beta f} + (z_{max} - \bar{z})\right) e^{\beta f (z_p - \bar{z})} \operatorname{erf}\left[\frac{\beta f \sigma}{2} + \frac{z_p - \bar{z}}{\sigma}\right] && \text{(5.6c)}\\
+ \; & (z_{max} - z_p)\left(\operatorname{erf}\left[\frac{z_p - \bar{z}}{\sigma}\right] - \operatorname{erf}\left[\frac{z_p - z}{\sigma}\right]\right) \;\Big\} && \text{(5.6d)}
\end{aligned}$$

with proportionality to $\frac{\sqrt{\pi}\sigma}{2k_B T}\Delta\xi$. Fixing $\bar{z} = 11$ Å, $z_f = 5$ Å and $z_{max} = 29$ Å, leaves Eq. (5.6) dependent on the friction peak position z_p, width σ and the force constant f. We abbreviate the result to

$$\Delta\mathcal{T} = \frac{\sqrt{\pi}\sigma}{2k_B T}\Delta\xi \cdot g(z_p, \sigma, f) \tag{5.7}$$

which is the product of the friction peak height, width, and a scaling factor $g(z_p, \sigma, f)$. For the moment, we neglect the dependence of $g(z_p, f)$ on σ because the direct proportionality of $\Delta\mathcal{T} \propto \sigma$ is the dominating peak width dependence for our values of $\sigma \approx 1$ Å. In some sense, the factor $g(z_p, f)$ quantifies the impact of the friction peak, which is why we refer to it as the *impact factor*.

If we choose the slope of $f = 13$ $k_B T$/nm from Fig. 5.5 (b.1), the impact factor steeply increases with increasing peak position of the kinetic profile as shown in Fig. 5.6 (b). Moreover, the broken line types indicate how $g(z_p, f)$ increases with decreasing potential slope f whereas the thick black line is the case for $f = 0$. If the force constant even becomes repulsive, the impact factor increases even more. In that case, the repulsive potential slope and the repulsive kinetic barrier add up. As a conclusion, we find that the mean binding time is proportional to the friction peak height $\Delta\xi$. Though, the impact can drastically decrease if the peak shifts downward an attractive slope. For repulsive slopes, the impact can drastically dominate such that, for our case, the unbinding should be dominantly affected.

5.4 Unbinding

Repeating the previous calculation for unbinding, i.e., starting the process at z_f to escape to a target position outside the pocket, even yields a scaling proportional to $g \propto \exp(\beta f(\bar{z} - z_f))$. Hence, the unbinding from our pocket is even more dominantly affected by the slope of the PMF and the potential depth. So for unbinding, the impact

factor takes values much larger than one whereas, for the binding, it takes values smaller than one. Here, we want to compare the impact on average binding and unbinding times. Therefore, we employ Eq. (2.14) from Chapter 2 for two scenarios – one neglecting the kinetic barrier, namely $\xi(z) = \xi_\infty$, and one incorporating the kinetic barrier, thus $\xi(z) = \xi^{\mathrm{M}}(z)$.

For the binding MFPT curves, we choose the integration boundaries in Eq. (2.14) such that the process starts at z_{max} and ends at the target distance $z_{\mathrm{b}} = 1$ nm. The average binding time is then calculated as the average over the full spatial range by

$$\mathcal{T}_{\mathrm{on}} = \frac{\int_{z_{\mathrm{b}}}^{z_{\mathrm{max}}} \mathrm{d}z' \mathcal{T}(z', z_{\mathrm{b}})}{(z_{\mathrm{max}} - z_{\mathrm{b}})} \,. \tag{5.8}$$

The average binding times $\mathcal{T}_{\mathrm{on}}$ *with* and *without* kinetic barrier are listed in Tab. 5.1. The variation in the absolute values upon the different ligands only arises due to the different bulk friction constants (see Appendix D.1). This impact of the the bulk friction is clarified in the histogram of Fig. 5.6 (a.1), which plots the tabulated $\mathcal{T}_{\mathrm{on}}$ values normalized by the respective bulk friction value. These normalized values are comparable and the kinetic barrier adds roughly an extra 50—70%. This relative deceleration can also be observed as we plot the ratio of $\mathcal{T}_{\mathrm{on}}$ *with* and *without* kinetic barrier in panel (b.1). Hence, the overall impact of the kinetic barrier on the binding times only yields a small increasing factor between one and two.

For the unbinding MFPT curve we choose in Eq. (2.14) the process to start at $z_{\mathrm{f}} = 0.4$ nm and to reach $z_{\mathrm{ub}} = 1.2$ nm. The average is equivalently evaluated inside the MFPT curve's full spatial range such that

$$\mathcal{T}_{\mathrm{off}} = \frac{\int_{z_{\mathrm{f}}}^{z_{\mathrm{ub}}} \mathrm{d}z' \mathcal{T}(z', z_{\mathrm{ub}})}{(z_{\mathrm{ub}} - z_{\mathrm{f}})} \,. \tag{5.9}$$

In contrast to the average binding times, the average unbinding times with and without kinetic barrier are dominantly affected as we see in Tab. 5.1. We see that $\Delta\mathcal{T}_{\mathrm{off}}$ can reach up to hundreds of microseconds, which arise from contributions

Tab. 5.1: Average binding times $\mathcal{T}_{\mathrm{on}}$, average unbinding times $\mathcal{T}_{\mathrm{off}}$ with and without kinetic barrier, respectively, as well as their differences for all aromatic ligands

kinetic barrier	$\mathcal{T}_{\mathrm{on}}$ [ns] without	$\mathcal{T}_{\mathrm{off}}$ [µs] without	$\mathcal{T}_{\mathrm{on}}$ [ns] with	$\mathcal{T}_{\mathrm{off}}$ [µs] with	$\Delta\mathcal{T}_{\mathrm{on}}$[ns]	$\Delta\mathcal{T}_{\mathrm{off}}$[µs]
ethylbenzene	0.943	49.0	1.613	270	0.670	221
toluene	0.926	12.0	1.511	44.4	0.585	32.4
benzene	0.868	2.56	1.243	5.20	0.375	2.64
phenol	0.828	0.465	1.210	1.12	0.382	0.655
benzyl alcohol	1.039	2.90	1.453	6.56	0.414	3.66

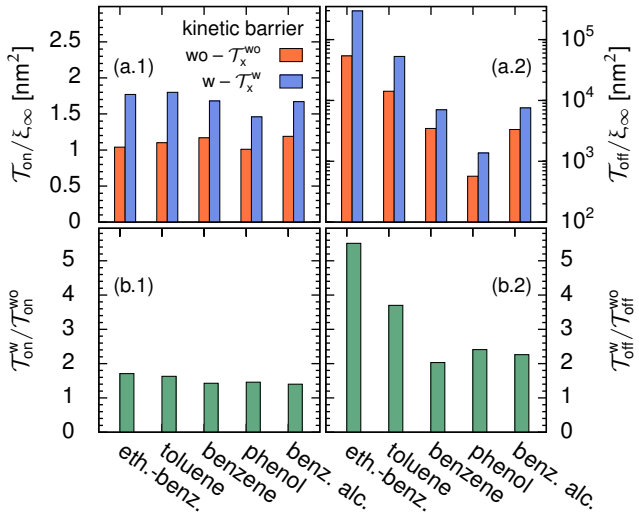

Fig. 5.7: Panels (a.1) and (a.2) show histograms of the average binding time and the average unbinding time normalized by the respective ligand's bulk friction value ξ_∞. The values with kinetic barrier (blue) are clearly enhanced in comparison to the values without kinetic barrier (orange). (b.1) The ratio of the average binding time with and without kinetic barrier exhibits a constant impact upon the various ligands. (b.2) In contrast, if the ratio of the average unbinding times with and without kinetic barrier is plotted it is obvious that the unbinding times dominantly increase by a factor of more than five for ethylbenzene.

of the PMF in $g \propto \exp(\beta f(\bar{z} - z_f))$. Moreover, the ratio of the average unbinding times *with* and *without* kinetic barrier in Fig. 5.7 (b.2) yields a factor of more than five for ethylbenzene and is generally non-constant for our various ligands. Note, however, that the proper kinetic barrier for the unbinding process can differ from the $\xi^M(z)$-profiles which we originally extracted from the binding process. This process is most sufficient and efficient for the conclusive interpretation of our estimates of the unbinding times. Thus, we assume that the conclusions and implications remain the same since, in particular, the energy landscape, or binding affinity, mainly steers the impact (factor) of a kinetic barrier height $\Delta\xi$.

On the one hand, the resulting estimates for the unbinding time confirm that more (extended) hydrophobic ligands reside longer inside the pocket. On the other hand, if we neglect the additional dissipative forces from the kinetic coupling the unbinding estimate can be wrong by several hundred microseconds and a factor of more than five. In particular, ethylbenzene would stay for 270 µs, if determining the binding time from data that incorporate the kinetic barrier, while it would only stay for 49 µs if we ignore the kinetic barrier. In summary, we find a constant impact on the binding times, while the impact on unbinding times predominantly changes.

5.5 Concluding remarks

In this final chapter, we presented the implications for an additional degree of freedom – a ligand's orientation. Taking our non-spherical aromatic ligands solvation effects seemed to steer the orientation close to the pocket. All ligands tended to desolvate their molecular components which were most hydrophobic. This behavior was especially pronounced for, e.g., phenol and benzyl alcohol, which both contain a hydroxyl group. For other ligands, this led to two optimal orientations. In particular, the rotation of toluene and ethylbenzene represented an additional bimodally fluctuating degree of freedom. Moreover, we found that more wetted binding sites would more heavily penalize the reorientation pathway. Hence, more dewetted binding sites, such as our pocket, would loosen the reorientation pathway.

Finally, this chapter served as our conclusion about the non-Markovian kinetics and Markovian methodology. From previous parts of this thesis, we knew that the kinetic coupling between pocket hydration fluctuations and enhanced ligand friction was proportionally related (Chapter 3). And we knew that the non-Markovian retardation manifested spatially shifted (Chapter 4). Here, we elaborated on a rescaling procedure which, by construction, reproduced the correct kinetics. The rescaling reinterpreted the enhanced dissipative forces as a kinetic barrier, whereas we also quantified the impact of the kinetic barrier depending on its position and relative to the attraction in the PMF. As a key result, we found that the kinetic coupling majorly affected unbinding rates rather than binding rates. We reached factors of more than five in the unbinding times, while the binding was slowed by less than a factor of two. After all, the rescaled results were conclusive; however, we employed a Markovian framework.

6 Final remarks – learning from modeling

It has become increasingly appreciated, that the optimization of active compounds for their overall efficacy in open, for instance *in vivo*, systems must reach beyond a simple maximization of the affinity to a respective receptor. One important aspect, increasingly amenable to *in silico* studies, is the kinetics of assembly with the target (macro)molecule. In this respect, hydrophobicity-mediated interactions, with their wide variety of underlying physical effects, seem to offer appealing possibilities for controlled adjustments for instance within the scope of synthetic cavitands or drug discovery.

In this work, we tackled such effects of hydrophobicity on binding kinetics and, thus, extended the discourse beyond the well-studied implications for binding thermodynamics. In particular, we based our investigations on the original work from Setny *et al.* [55], who first found the coupling of ligand kinetics and a binding site's hydration fluctuations. Our primary incentive was to rationalize this *kinetic coupling* on a quantitative level and to understand how it could thus serve for smart drug development.

6.1 Summary

In our first investigation in chapter 3, we started from a most minimalistic point of view. We developed a simple stochastic model of two nonlinearly coupled equations to map the two fundamental features of water-mediated key-lock association. The first feature was the *bimodally fluctuating pocket hydration*. The second feature was the well-known result that the solvation free energy of a nanoscopic hydrophobic ligand scales with the solvent-excluded volume [7] which we implemented as the coupling potential – the *desolvation potential*. On the one hand, this coupling potential drove ligand association and, on the other hand, it incorporated the required capillary drying inside the pocket. The full two-coordinate model enabled us to tune pocket hydration by biasing the wet or dry state and modifying the time scales of water interface fluctuations, separately. These modulations could result from more or less hydrophobically confining pockets; however, realistic pocket modifications non-trivially tune hydration biasing and fluctuations at the same time.

We analyzed the FPT problem for ligand binding in this model, which revealed non-Markovian contributions that were consistent with Setny's original work [55]. The numerical solution of the two-dimensional model showed decelerated binding in comparison to an effective one-dimensional theoretical description which employed the equilibrium PMF and a spatially constant friction. Resolving the latter, we found that it was spatially dependent and maximized at the position where the potential coupling between the ligand and the water interface emerged. Most strikingly, if we increased the time scale of the water interface fluctuations, the ligand PMFs remained unchanged, however, the deceleration of ligand binding was enhanced. This result made evident that the ligand kinetics coupled to the time scale of water fluctuations. From these findings, we interpreted that the additional force fluctuations induced the enhanced local friction that arose from the coupling potential between the ligand and the water interface, i.e., the desolvation potential.

We further corroborated our novel interpretation by constructing a generalized Langevin model, i.e., a one-dimensional GLE. The GLE employed the second order perturbation of the desolvation potential as damping strength and applied the water interface's auto-correlation time scale in an exponentially decaying memory kernel. As one of our core results, we found good agreement between the friction of the one-dimensional GLE and the friction in the previous two-dimensional model, i.e., the two coupled equations. Moreover, the fluctuation-dissipation relation of the GLE expressed a proportional scaling of the friction peak with the fluctuation time scale of the water interface. In summary, our GLE substantiated that non-Markovian kinetics within the coupled stochastic system – and seemingly within MD simulations – was the outcome of projecting the binding kinetics onto the ligand reaction coordinate alone.

As a next step, in chapter 4, we wanted to understand the actual, convoluted tunability of pocket hydration fluctuations in explicit-water MD simulations. We principally tackled Ball's notion of *modulating the water properties regarding chemical and geometric features* [100] of our pocket when we systematically modified these physicochemical properties of the binding site. Thus, this approach considered a pocket's general degree of hydrophobic confinement. The spectrum of our modifications in geometry and water-wall attraction strength was likely to reflect naturally occurring variations in shape and packing density of receptor binding sites. Also, the properties of synthetic hosts, e.g., synthetic cavitands, should be adjustable within our considered range of pocket modifications [60, 62].

Our general, intuitively expected finding was that the binding rate positively correlated with an increasing degree of hydrophobicity. However, our results less intuitively highlighted that the binding kinetics were not directly governed by host-guest interactions but rather indirectly, substantially mediated by the water effects. Specifically, we pointed out that the degree of pocket hydration and the amplitude of solvent density fluctuations could be considered as standalone descriptors of the binding times. Such a dependence on solvent behavior led to drastic effects in some cases. For example, increasing the pocket depth merely by one ångström was enough to switch from slow to fast binding because the magnitude of solvent fluctuations abruptly increased.

Subsequently, we further dissected the maximized ligand friction prior to binding. We employed the fluctuation-dissipation relation, which we previously derived from the GLE description in chapter 3. This relation well applied to the bulk of our investigated physicochemical pocket variations. Thus, we could show that the slowing ligand friction in the MD simulations was also directly tuned using the hydration fluctuation time scale, which could be a potential obstacle that must be overcome if binding rates should be optimized. Therefore, we interpreted our fluctuation-dissipation relation of the kinetic coupling as a system specific rationalization of Philip Ball's proclamation to unravel *how the dynamics of biomolecular solutes relate to those of water and how fluctuations would couple* [100].

Finally, we laid out that quantitative binding time predictions must account for the time transients of the solvent response. Technically, the fluctuations materialized spatially shifted if an umbrella potential was or was not constraining the ligand. Thus, if we probed the friction under ligand restraining conditions, we had to incorporate a delaying spatial shift in order to improve the theoretical binding time estimates. Future work along these lines should show for what kind of realistic scenarios these effects could be important.

In the final chapter 5, we put our focus on non-spherical ligands, which thus added another degree of freedom – rotation. Generally, the pathway of ligand orientation upon binding was seemingly steered by solvation free energy. If the investigated ligands were in the vicinity of the binding site, they favored certain orientations such that they could partially desolvate their most hydrophobic sections. Moreover, binding to the dewetted pocket comprised two advantageous over binding to a more wetted binding site. First, the "smeared", i.e., dewetted, water interface in the pocket weakened the energetic stabilization of the orientation pathway, and second, it stretched the whole reorientation process over a wider spatial range

before association. Specifically, if a ligand favored two distinct alignments, the rotation represented an additional bimodally fluctuating degree of freedom, which also coupled to the ligand kinetics, i.e., enhancing the local friction peak.

We finally discussed the non-Markovian effects again by asking how the kinetic coupling influenced the binding and the *unbinding* times. Therefore, we returned to the Markovian method from Eq. (2.15) to extract the *kinetic profile*, which captured the peak of the – generally speaking – dissipative forces. We re-interpreted this peak as a barrier, which was scaled into a new energy landscape using a method based on the Fokker-Planck equation [146, 147]. We rationalized the effect of the kinetic barrier regarding an *impact factor*, which for instance depended on the peak position relative to the attractive slope in the original PMF. More importantly, the impact factor was non-linearly decreasing with increasingly attractive slopes, and it consistently increased if the barrier would add to an anyhow repulsive slope. Hence, we found that the kinetic barrier only minorly influenced the binding times; however, it majorly decelerated unbinding. Our most hydrophobic ligands most pronouncedly obeyed this result, which confirmed that more hydrophobic ligands resided longer inside a hydrophobic binding site. Nevertheless, the fluctuations of the pocket hydration and ligand orientation substantially mediated this effect in contrast to the usual thermodynamic interpretation of enhanced binding affinity with increased hydrophobicity.

6.2 Discussion

To put it in a nutshell, we illuminated the non-Markovian kinetics of hydrophobic key-lock binding far from equilibrium by means of the theory of stochastic processes and newly extended explicit-water MD simulations. We discussed the non-Markovian effects using existing Markovian methods when we interpreted the additional dissipative forces as kinetic barriers. In Eq. (4.1), we rationalized that the kinetic barriers were proportionally related to the water fluctuation time scale. In the light of the FDT, this equation represented our system specific fluctuation-dissipation relation. Having derived this relationship for the bimodal hydration, we found that other bimodally fluctuating degrees of freedom, such as ligand orientation, also contributed to the kinetic barrier. Further, we explicitly calculated the slow down of the binding time in Eq. (5.7). This expression clarified that the barrier mainly affected the rates if the attractive slope of the PMF was weak or even repulsive. Thus, the kinetic barrier most importantly influenced the unbinding process in our systems.

To highlight this important notion again, we explicitly depict in the following discussion how our rationalization of the kinetic barriers in principle enhances a drug's residence time. Taking the fluctuation-dissipation relation in Eq. (4.1) with the slow down in Eq. (5.7), we can rewrite

$$\Delta\mathcal{T} = \frac{\sqrt{\pi}\sigma}{2}\tau_N \frac{\partial^2 V(z)}{\partial z^2}\bigg|_{z_p} g(z_p, f)\,. \tag{6.1}$$

This expression linearly scales with the time scale τ_N, which indeed suggests a linear tunability via the fluctuations. The second spatial derivative of the PMF also linearly influences the slow down, which also certainly infers a linear tunability. More importantly, the slope f of the PMF non-linearly tunes the impact $g(z_p, f)$ of the kinetic barrier, which infers *non-linear* tunability. Therefore, we suggest that, after all, the energy landscape and, thus, the binding affinity most dominantly steers the kinetic barrier.[1] For instance, increasing the binding affinity in a simple example deepens the energy well, which steepens the slope. Therefore, the impact of a kinetic barrier on unbinding times is larger for high affinities. Consistently, the impact on binding times narrows for high affinities.

This general thinking is still simple but highlights that the kinetic barrier is a substantial target for future investigations. Residence times proofed to be of fundamental importance because they positively correlate with a drug's efficacy [96–99]. Hence, we suggest the kinetic barriers as a principal target in future developments for control of a toxin's or drug's kinetics. In the case of a toxin, a kinetic barrier and its impact should be minimized, while for drugs both should be exploited to a required maximum. Certainly, future investigations along these lines have to find out which kinetic barriers are practically accessible. Notably, the resolution of hydration water dynamics progressively improves thanks to cutting-edge experimental developments [162–164].

Furthermore, we showed that any fluctuating degree of freedom could potentially slow unbinding rates when we resolved the ligands' rotation. Therefore, specific applications determine whether it is more important to steer a binding site's hydration or to modify other internal degrees of freedom. Interesting examples can be found in natural sciences as well as chemical engineering, such as the association of biomolecules with fluctuating charge or dipole moment [165, 166], voltage gated transport through membrane proteins [167], conformationally gated ligand binding to proteins [168–177], polymer translocation [178] or folding [179], and catalysis in

[1] The energy landscape, here the PMF, is directly linked to the binding affinity: enhancing the energy well in the PMF increases the binding affinity.

stimuli-responsive nanoreactors [180–183]. In specific cases, coupling fluctuations can yield optimal binding rates, in a “resonant” coupling regime [182, 184]. This fact poses the immediate question whether hydration fluctuations in a binding site can also couple resonantly to a ligand’s binding rate. Our stochastic model could enable initial investigations in this direction.

As a final critical notion, we add that our model system was restricted to its one-dimensional treatment. In particular, Tiwary *et al.* [104] critically assessed the one-dimensional restraint in a similar MD setup, where he found that the ligand least likely enters via a pathway that would include enhanced water fluctuations. Nevertheless, more realistic, i.e., complex, systems could intrinsically steer pathways for which drying transitions still play a role. Oppositely, we now know that water fluctuations can guide possible pathways. A ligand might particularly avoid a route which comprises kinetic barriers. After all, the presence of such drying transitions suggests that nature can utilize their effect if required. This relevance can also extend to fundamental folding and function of proteins, where dewetting plays a role [185] while folding is the process that hides most hydrophobic sections of a protein into its interior [67, 69]. The kinetics in protein folding have previously been explored by Hinczewski *et al.* [146] when they introduced the aforementioned rescaling procedure. One of their main conclusions was that the importance of novel features in the rescaled energy landscape especially increased due to explicit water effects which introduced new kinetic mechanisms.

6.3 Outlook

In the following, we want to discuss key aspects that remain open to future work: the FPT problem for spatially dependent non-Markovian kinetics, a capillary continuum description of the water interface at surfaces with various physicochemical properties, and the next step to search for real binding sites that suit the conditions of our model.

For the first problem, we want to present one idea that could address the FPT of non-Markovian kinetics with a coordinate dependent memory kernel; however, other approaches could work [186]. To the knowledge of the author, current theories lack such a description of a process in which the time transients of the memory kernel depend on the reaction coordinate itself. For such a process, even the fulfillment of the FDT would have to be questioned and worked out.[2]

[2]Note also that, e.g., spatio-temporal kernel functions differ to the spatially dependent non-Markovian effects we encountered in this thesis.

In particular, we suggest a first solution by the adjoint operator approach from Eq. (2.13), which at first place is only valid for a Markovian process. However, sticking to this methodology a non-Markovian formulation is principally possible. In that case, the approach needs to adjust the adjoint operator to prevent back-flow, namely, trajectory contributions that might have left or even originated outside the considered spatial range from which the process escapes [141]. Explicit adjustments are difficult and have been solved for "simplest" non-Markovian processes such as telegraphic noise [142]. Exact adjustments are particularly rare. Nevertheless, approximate methods still exist and could be applied [187]. Such methods help to formulate an evolution equation [188] of the underlying non-Markovian process via, e.g., an expansion in the cumulants [187, 189]. In simple cases, the resulting operator $\Omega^\dagger$ might ascribe an integro-differential equation that is essentially memory-free. After all, the resulting integro-differential equation of the type of Eq. (2.13) is left to be solved.

An important note is that such non-Markovian theories are even desirable for applications far from (bio)chemically relevant systems. The "minimal" requirement for a theory of this type is a kinetic coupling that yields local time transients because of locally coupled degrees of freedom. The experienced reader is certainly aware that the stochastic dynamics on multidimensional energy landscapes produce an actually local friction and, possibly, locally dependent memory effects in a reduced, lower dimensional description [190–192]. As mentioned before, in the case of bimodal hydration fluctuations in a ligand-pocket system, a two-state approach has been successfully applied [101]. This particular example counts to the systems for which the dynamic components are comfortably accessible; however, this accessibility is in many cases far from obvious. For these cases, we also suggest to reside to advanced sampling methods that are being developed for increasing simplicity and efficiency throughout an ongoing effort of contemporary research [104, 105, 173, 193, 194].

The second issue would be a capillary theory of solvation that can predict water dynamics at heterogeneous interfaces such as proteins and membranes. In this thesis, we measured the degree of hydrophobic confinement via the hydration fluctuations. Predicting the water dynamics via a coarsened theoretical approach could be desirable. For water, first approaches go back to the theory of Lum, Chandler, and Weeks [14]. They reduced the picture of water at solutes to its density fluctuations regarding time-independent correlation functions which, i.e., quantify hydrophobicity. Their work set a cornerstone for a solid framework of the solvation thermodynamics and wetting [44, 195], and also initially justified

our phenomenological measurements of χ_N. Without further detail, one approach could extend these theories to time-dependent correlation functions and, thus, time-dependent linear response theory of capillary water surfaces.

Further, a consistent formalization of dynamic phenomena of capillary water interfaces is a current and future objective of capillary wave theory. For these theories, a surface's undulation is an energetic penalty in an Hamiltonian formulation of the interface. Theoretical as well as experimental studies investigated first models for such capillary surface waves at planar surfaces [196–198]; however, the dynamic response of capillary waves at arbitrary solute interface properties also remains a challenge to future work. In principle, a dynamic model of water on such a continuum level of solvation could help for extended simulation studies by more accurate implicit solvent models [199–201].

Finally, we suggest the immediate search for macromolecules, such as proteins, that explicitly exhibit drying transitions, i.e., hydration fluctuations, inside a binding site. Primary motifs can be any cleft without particular function as a binding site, however, as a final goal the study should focus on finding a specific ligand-receptor in which hydration fluctuations are relevant. Particularly, we propose to start with the streptavidin-biotin complex, the Cox-2-arachidonic acid complex, the antibody DB3-aetiocholanone complex, and the HIV protease receptor, which have been shown to contain several more or less-pronounced hydrophobic enclosures.[3] These choices are based on a study in Ref. [56], which discussed the thermodynamics of binding into these pockets. Also in Ref [87] more examples have been depicted and characterized concerning the drying transitions in a given protein's active sites. We also encourage future studies that exploit further simple, internal degrees of freedom that might yield kinetic couplings. Particularly, the binding of the anti-cancer drug Dasatinib to its src kinase receptor was shown to exhibit fluctuating mean forces, which also rooted from conformational fluctuations of the receptor [106]. The kinetic modeling of this system required additional barriers accounting for the reorganization of the receptor; however, the water fluctuations of the receptor seemed to be even more relevant.

[3]The HIV protease receptor has actually no hydrophobically enclosed sections and could thus serve as reference.

Appendix

A – Setny pocket

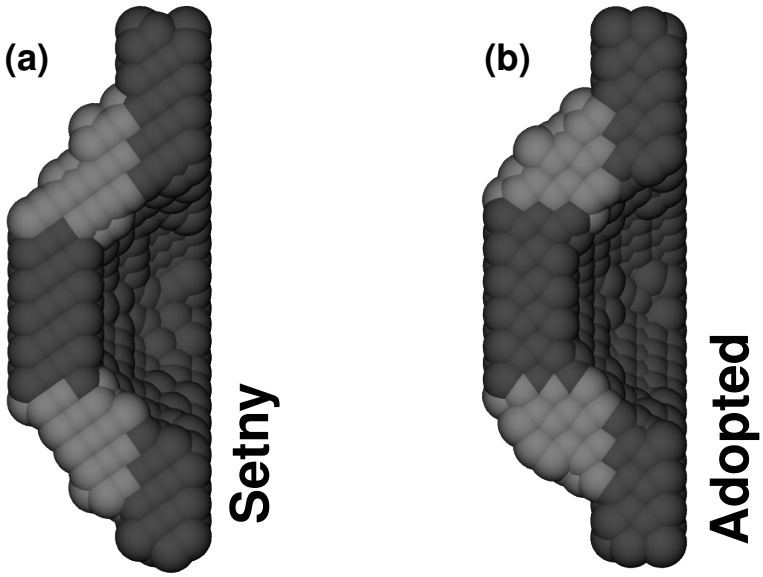

Fig. A.1: Comparison of (a) the original model pocket by Setny [48–50] with (b) our model pocket constructed by an adopted procedure which was presented in Section 2.2.3. The pocket edges are highlighted by silver beads to illustrate the different thickness in both models. The extra beads on the back of our wall in (b) add to an overall higher LJ attraction. Note also that the crystal structures slightly differ, as it is visible if one looks at the details of the bead ordering. We utilize the HCP crystal structure described in Section 2.2.3.

For the biggest part, we motivate our setup by the one from Setny [48–50]. We adopted his model of a hemispherical surface recess in a hydrophobic wall. In our case, we construct the pocket using a spherical probe volume to consistently build various pocket geometries (regarding curvature and depth) as we presented in Section 2.2.3. To compare the original pocket from Setny with one of ours, we show in Fig. A.1 two screenshots. Silver beads highlight the pockets' edges. These edges are slightly thicker in the pocket, which was constructed by our adopted procedure. Hence, an extra ring of LJ beads at the back of the pocket is additionally attracting water into the pocket. Remember that the crystal layer thickness is $c/2 = 1.02$ Å

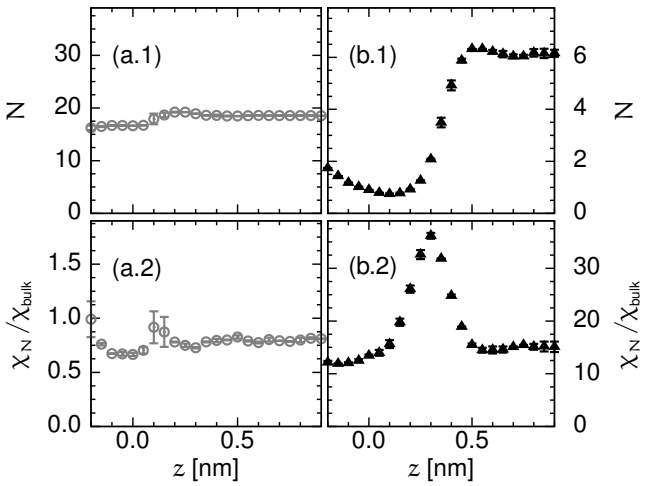

Fig. A.2: If we use the original LJ parameters [55] with our adopted pocket from Fig. A.1 (b) the pocket water occupancy in (a.1) is almost 20 and the normalized fluctuations in (a.2) are even slightly smaller than in bulk. If we use our adopted dispersion attraction parameter $\epsilon_p = 2.4$ J mol^{-1} the occupancy decreases with ligand distance and the normalized fluctuations peak for intermediate ligand-pocket separations, which best reproduces the hydration fluctuations from Ref. [55].

such that this extra ring of beads lies well within the cutoff range of one nanometer in the simulations.

We compensate the extra LJ contributions by our smaller dispersion attraction, i.e., our smaller LJ parameter ϵ_p for the wall particles. In our case the value reads $\epsilon_p = 2.4$ J mol^{-1} where in the original publications from Setny it was 39.3 J mol^{-1}. As a result, we obtain similar hydration fluctuations, although, the extra ring of beads adds to the overall water-wall attraction.

For a better comprehension of this remarkable difference between the original parameters and ours, we compare in Fig. A.2 the average number N of water molecules inside the pocket and its normalized fluctuations χ_N/χ_{bulk}. Here, we employ our setup and compare simulation results of our weak parameter settings to the results of the originally stronger LJ attraction. Using the original dispersion attraction strength in our pocket yields an extremely increased pocket wetting with much weaker fluctuations in Fig. A.2 (a.1) and (a.2). Using our smaller parameter in panels (b.1) and (b.2) results in the low wetting and peaking fluctuations that are crucial for the kinetic coupling of the ligand and the pocket water. Most importantly, the new parameterization and the results in Fig. A.2 (b.1) and (b.2) best reproduce the hydration fluctuations from Ref. [55].

B.1 – System size dependence of MFPT and $\xi^M(z)$

In general, the peak height of $\xi^M(z)$ can easily be shown to be system size dependent due to the way Eq. (2.14) and thus Eq. (2.15) are defined in Section 2.3. The MFPT

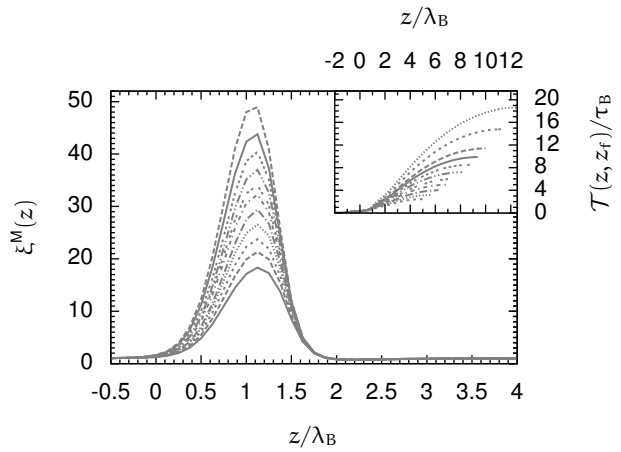

Fig. B.1: The peaks in $\xi^M(z)$ increase with system size, namely, if the reflective boundary z_{max} outside the pocket is further away. The peaking values are proportional to the slopes in the MFPT curves, shown in the *inset*, which increase with z_{max} due to extendedly available trajectory paths for the ligand migration. The peaks of Eq. (2.15) in the main plot thus increase with increasing z_{max}.

from a point z to cross z_f depends on the choice of the reflective boundary z_{max}. This becomes most evident if one considers for example a process with constant energy landscape $V(z) = 0$ and constant friction Eq. (2.14) then simply yields

$$\mathcal{T}(z, z_f) = -\frac{1}{2}\beta\xi \left[(z^2 - z_f^2) + z_{max}(z - z_f)\right] .$$

Thus, the MFPT at each position z linearly increases with z_{max}, which contributes to an increasing slope of the curve which in turn enters Eq. (2.15).

Certainly, this behavior can also be observed from numerical results of our coupled stochastic model in Eq. (3.6) from Chapter 3. The MFPT curves shown in the inset of Fig. B.1 increase with z_{max} due to extendedly available trajectory paths. Thus the slopes of the curves increase. Since the profiles $\xi^M(z)$ from Eq. (2.15) are proportional to the derivative $dT(z, z_f)/dz$, their peaking values are enhanced with z_{max}, which is illustrated in the main plot of Fig. B.1.

B.2 – Umbrella sampling for the stochastic model

In order to calculate the friction $\xi(z_i^r)$ by means of the PACFs via Eq. (2.11), the umbrella setups are restraining the ligand at positions $z_i^r/\lambda_B = 1$ to 2.625 in steps of 0.125. The choice of the interval is made by an initially coarse scan from positions deep inside the pocket up to distances far away. The resulting values $\xi(z_i^r)$ are fitted by a Gaussian, which is illustrated in Fig. B.2 for setups which utilize barrier heights from $h = 1\ k_BT$ to $5\ k_BT$ in steps of $0.5\ k_BT$. A second friction peak is also present inside the pocket around $z/\lambda_B = -1.5$ as it is plotted in the inset of Fig. B.2. Its dependence on double well barrier height is similar because the essential underlying

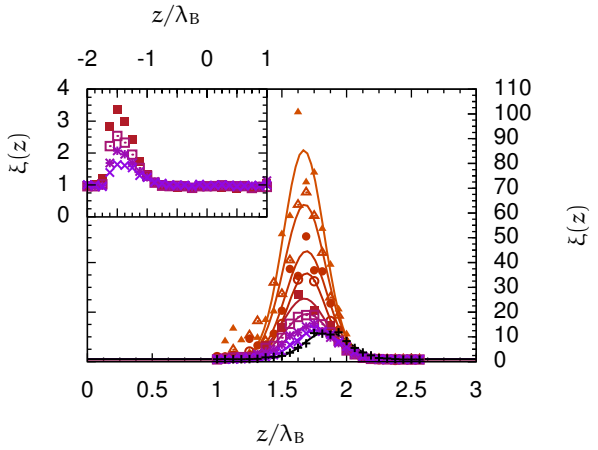

Fig. B.2: The symbols represent spatial friction coefficients $\xi(z_r)$ from PACFs via Eq. (2.11). The peaking values increase with increasing double-well barrier height represented in varying color. Gaussian fits to the respective data are plotted as solid lines. The *inset* shows similarly peaking friction when the ligand is inside the pocket and thus randomly interacts with the interface, whereas here $h = (1.5,\ 2,\ 2.5,\ 3)\ k_BT$. The color coding is adopted from Fig. 3.3 in Chapter 3.

reason is the same but it is not of further relevance to our discussion. A doubled spring constant gave similar values for $\xi(z_i^r)$ within errors of approximately 5%, thus, confirming a sufficient choice of the spring constant.

Note also that sampling has to be increased if the barrier height is increased in order to sufficiently sample slowed water fluctuations. Hence, elongated trajectories are calculated for statistically converged PACF calculations; however, the data remains more noisy for simulations with extended water interface fluctuation time scales.

B.3 – Separation of the ligand and water interface

In order to develop our GLE model in Section 3.3, we expand the coupling desolvation potential up to second order in $\zeta(z_s, z_l)$. It is used to identify the damping constant $\Omega = 2\pi\Gamma\epsilon \equiv \xi/\tau$ in Eq. (3.14). For comparison to the peaking friction values from the PACFs, we extract the average and, thus, dominant distance ϵ between the water interface and the ligand. Therefore, the ligand is fixed by an umbrella potential at the positions of each friction peak to obtain the distributions $P_h(\zeta)$ of the ligand distance to the water interface. The distributions depend on h and are plotted in Fig. B.3. The average $\epsilon(h) = \langle\zeta\rangle$ and the standard deviation $\Delta\epsilon(h) = \langle\zeta^2 - \langle\zeta\rangle^2\rangle^{1/2}$ are calculated within the interval $[0, 1]$, in which the interface and the ligand interact. The values $\epsilon(h)$ are plotted in the inset of Fig. B.3 with their standard deviation as error bars. The black line draws the mean $\epsilon = \langle\epsilon(h)\rangle_h = 0.36$, which we used in the coupling strength Ω in Section 3.3. Similarly, the average standard deviation $\Delta\epsilon = \langle\Delta\epsilon(h)\rangle_h = 0.24$ is plotted as blueish shade in Fig. B.3.

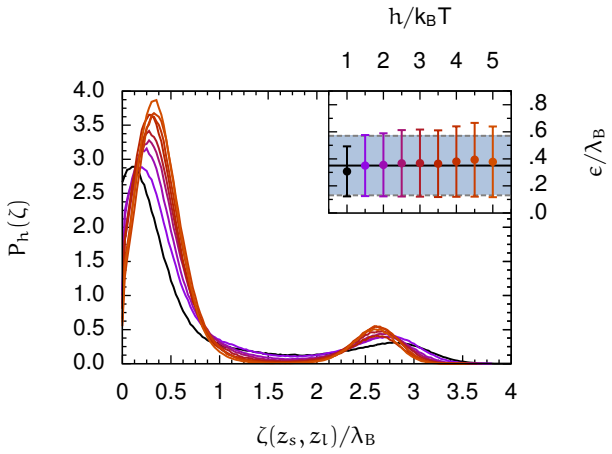

Fig. B.3: The distributions $P_h(\zeta)$ are drawn for varying double well barriers while the ligand is restrained at the peaking position of the friction from Fig. 3.7. The *inset* plots the respective first moment with the second moment as error bars evaluated from $P_h(\zeta)$ in the interval $[0, 1]$. The average of the first moment is drawn as black line and the blueish shade illustrates the average of the second moment. The color coding is adopted from Fig. 3.3 in Chapter 3.

C – Ligand PMF, friction and water fluctuations

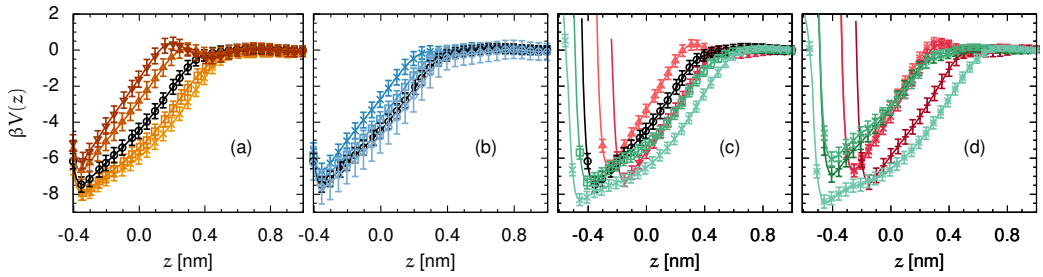

Fig. C.1: The PMF $V(z)$ changes if (a) the pocket-water attraction ϵ_{pw}, (b) the pocket radius r_i or (c) & (d) the depth Δ change. In panel (a) it is evident that with increasing pocket-water attraction the PMFs manifest a barrier. In panel (b) the PMFs are only slightly shifted inwards if the pocket radius slightly increases. While the radius $r_i = 9.5$ Å is fixed in panel (c), differently deep pockets naturally shift the inner repulsion and exhibit a barrier in shallow pockets. Similarly, panel (d) shows the data for differently deep pockets with fixed $r_i = 9$ Å (red) and 11 Å (green). The symbol and color coding is adopted from Fig. 4.2 in Chapter 4.

In Chapter 4, we only use the essential features of the PMFs, the ligand friction, the pocket occupancy fluctuations, and their time scales to present a concise picture of the physics. For completeness, the full profiles of these observables are summarized here.

Fig. C.1 plots the PMFs $V(z)$ from $z = -0.4$ nm to 1 nm. The symbols and color coding are adopted from Fig. 4.2 in Chapter 4. Panel (a) plots the PMFs of the systems with tuned water-wall dispersion interaction. If hydrophilicity increases, the dominant attraction shifts into the pocket and a small barrier occurs. In panel (b), it becomes evident that the curvature minimally impacts the PMFs. Here, the

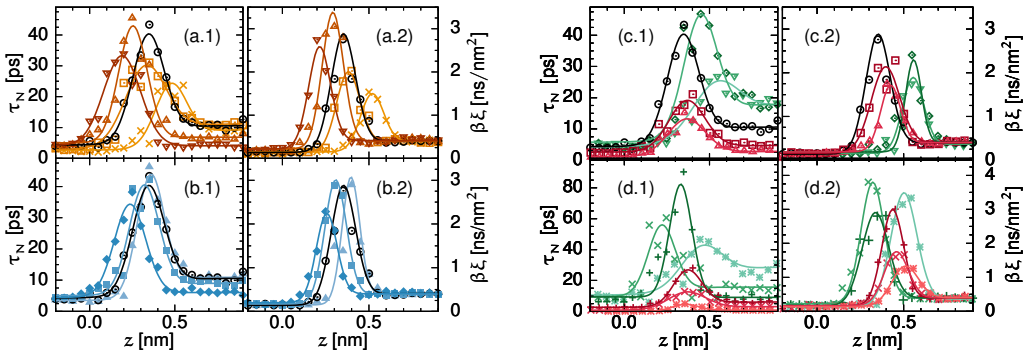

Fig. C.2: The maxima of $\tau_N(z)$ and $\xi(z)$ correlate and thus exhibit the kinetic coupling of Eq. (4.1) in Chapter 4. While the hydrophobicity decreases due to increasing water-wall interaction, the peaks of $\tau_N(z)$ in (a.1) and the peaks of $\xi(z)$ in (a.2) exhibit a similar non-monotonic trend. On the other hand, while curvature increases the common trend is monotonically decreasing, if one compares $\tau_N(z)$ in (b.1) and $\xi(z)$ in (b.2). In (c.1) and (c.2) both correlated measures show a non-monotonic behavior with pocket depth, while the radius is fixed to $r_i = 9.5$ Å. Panels (d.1) and (d.2) respectively plot the fluctuation time scale and ligand friction in the systems with radii $r_i = 9$ Å and 11 Å, while their depth is tuned. The colored lines are fits of Eq. (C.1). The symbol and color coding is adopted from Fig. 4.2 in Chapter 4.

dominant attraction also shifts slightly inwards with decreasing pocket curvature, i.e., increasing pocket radius. Changing the depth in panels (c) and (d), naturally shifts the inner repulsion of the PMFs and also the attracting slope. Notably, most shallow pockets also exhibit a small repulsive barrier similar to our most hydrophilic pockets in panel (a).

Fig C.2 shows the pocket occupancy fluctuation time scale $\tau_N(z)$ and ligand friction $\xi(z)$ profiles. The raw data is represented as colored symbols adopted from Fig. 4.2. The lines represent fits by

$$f(z) = l^2 + \frac{1}{2}(u^2 - l^2)\left(1 + \frac{e^{(x-\mu)/\sigma} - 1}{e^{(x-\mu)/\sigma} + 1}\right) + Ae^{-\frac{(x-\mu)^2}{\sigma^2}} \qquad \text{(C.1)}$$

with fit parameters (l, u, μ, σ) which are drawn with the color coding from Fig. 4.2. We pair the observables $\tau_N(z)$ and $\xi(z)$ because their peak values are coupled, which was depicted in Eq. (4.1) of Chapter 4. Here, this coupling also becomes evident in Fig. C.2. Panels (a.1) and (a.2) respectively plot $\tau_N(z)$ and $\xi(z)$ for the systems for which the hydrophobicity is tuned due to water-wall attraction. Panels (b.1) and (b.2) show the respective measures for varying pocket radius. The plots on the r.h.s. of Fig. C.2 draw the data for differently deep pockets with fixed radius $r_i = 9.5$ Å in panels (c.1) and (c.2), as well as for fixed $r_i = 9$ Å and 11 Å in panels (d.1) and (d.2). Comparing the various pockets, the timescale τ_N varies on two orders of magnitude, even if the ligand-pocket separation is large. Additionally, non-monotonic or monotonic trends with hydrophobicity or curvature in the peak

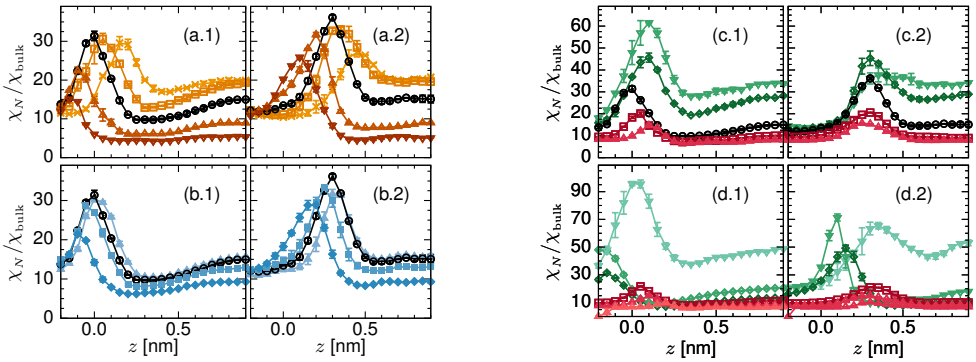

Fig. C.3: The pocket occupancy fluctuations from constrained and unconstrained simulations reveal the solvent response during the process of association. Panels (x.1) and (x.2) show $\chi_N(z)$ from unconstrained and ligand constraining simulations, respectively. Panels (a.1) and (a.2) draw $\chi_N(z)$ in the setups with varying hydrophobicity, i.e., water-wall interaction. The curvature sensitivity of χ_N is presented in panels (b.1) and (b.2). Panels (c.1) and (c.2) draw the data with fixed radius $r_i = 9.5$ Åbut varying depth Δ. Panels (d.1) and (d.2) show $\chi_N(z)$ when tuning the depth with radii $r_i = 9$ Å and 11 Å. The symbol and color coding is adopted from Fig. 4.2 in Chapter 4.

heights of the time scales transfer into the peak heights of the ligand friction. At the same time, the peaking positions consistently shift for both measures.

Fig C.3 plots the pocket occupancy fluctuation strength $\chi_N(z)$ in dependence of ligand-pocket separation. Panels (a.x) and (b.x) respectively plot the data with tuned hydrophobicity and curvature. Panels (c.x) and (d.x) show $\chi_N(z)$ for various pocket depths with fixed radii $R = 9.5$ Å and $R = 9.0$ Å, as well as 11 Å. In general, the fluctuations also maximize prior to ligand binding, which is associated with the capillary evaporation of the pocket. Though, the peaking structures from unconstrained simulations (in panels (x.1)) occur for smaller values of z than from ligand constraining simulations (in panels (x.2)).

The differently sampled occupancy fluctuation strength designates time transients, which differently manifest throughout the two simulation schemes. While the ligand is constrained, the system principally relaxes into an equilibrium situation, which is nevertheless highly dependent on z. The temporal transients of this relaxation are determined by τ_N. Since these are on the order of $\mathcal{O}(10\text{ ps})$, they are decaying during the association process in unconstrained simulations, while the ligand is moving along z. Thus, we note that the ligand binding occurs with a solvent response that cannot be captured in constraining simulations. In order to account for this situation during the analysis in Section 4.3, we shift the friction peaks from umbrella simulations by the respective shift in the data for χ_N. The shifted friction, then, can correct for some errors in the binding time estimation (see Fig. 4.5 in Chapter 4), however, we cannot correct for possible peak height or width variations in the shifted $\xi(z)$.

Tab. D.1: The different aromatic compounds have different bulk friction values $\beta\xi_\infty$. In general the more elongated/bigger ligands have larger friction constants.

	$\beta\xi_\infty$ [ns/nm^2]
ethylbenzene	0.91
toluene	0.84
benzene	0.74
phenol	0.82
benzyl alcohol	0.87

D.1 – Bulk friction constants of the aromatic ligands

Here, we briefly summarize the bulk friction constants of aromatic compounds. In general, we can probe the bulk friction in umbrella windows, that are far away from the binding site using PACFs and Eq. (2.11). For the spherical ligand, we know that the bulk friction is around $\beta\xi_\infty = 0.4$ ns nm^{-2}. According to Stokes friction, the bulk friction should increase with ligand size, which holds for our aromatic compounds. Table D.1 lists the bulk friction for ethylbenzene, toluene, benzene, phenol, and benzyl alcohol. Throughout Chapter 5, we use these values to normalize the mean binding time and the kinetic profiles. This enables us to compare the effective differences between our various ligands.

D.2 – MFPT for benzene binding to the wall

Taking the PMF for benzene binding to the planar binding site and the benzene bulk friction value, we can predict the mean binding time by Eq. (2.14). In principle, this prediction should hold for Markovian processes with spatially dependent friction but for benzene association to the wall the spatially constant value of $\beta\xi_\infty = 0.74$ ns nm^{-2} is sufficient. Fig. D.1 shows that the binding time from theory and simulation well coincide. Hence, we can easily predict the MFPT of benzene binding to the wall for any system size, i.e., arbitrary $z_{\max}$ in Eq. (2.14). Especially for comparison to the pocketed binding site in Fig. 5.4, we have to use the theoretical

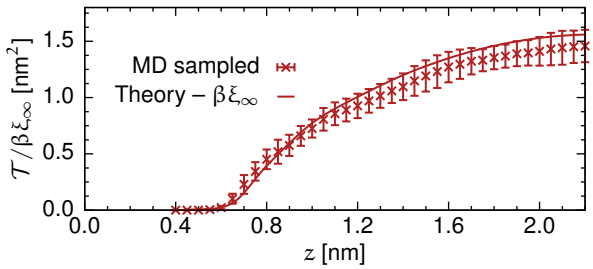

Fig. D.1: The simulation sampled MFPT curve (red crosses) for benzene association to the planar wall coincides with the results of Eq. (2.14) (red line).

prediction for binding to the planar site. Due to the recess in the wall, the reflective boundary is farther outside in the system with the molded binding site than in the one with the planar site. Fortunately, we can comfortably extend the binding time to large reflective boundaries by Eq. (2.14). In Fig. 5.4 (a.1) and (b.2), we particularly draw the theory for $z_{\mathrm{max}} = 2.9$ nm.

List of acronyms

AA all-atom

ACF autocorrelation function

DFT density functional theory

FDT fluctuation-dissipation theorem

FPT first passage time

GLE generalized Langevin equation

H hydrogen

H-bond hydrogen bond

HCP hexagonal close packed

LJ Lennard-Jones

M virtual interaction site

MD molecular dynamics

MFPT mean first passage time

O oxygen

PACF position autocorrelation function

PMF potential of mean force

UA united atom

vdW van der Waals

WHAM weighted histogram analysis method

List of Figures

List of Tables

Bibliography

[1]P. Ball. *Life's matrix: a biography of water*. Berkeley: University of California Press, 2001.

[2]F. Franks. *Water: a matrix of life*. Cambridge: Royal Society of Chemistry, 2000.

[3]B. Bagchi. "Water dynamics in the hydration layer around proteins and micelles". *Chem. Rev.* 105.9 (2005), pp. 3197–3219.

[4]F. Franks. *Water. A comprehensive treatise*. New York: Ed. Plenum, 1973.

[5]A. Luzar and D. Chandler. "Hydrogen-bond kinetics in liquid water". *Nature* 379.6560 (1996), p. 55.

[6]F. H. Stillinger. "Water revisited". *Science* 209.4455 (1980), pp. 451–457.

[7]D. Chandler. "Interfaces and the driving force of hydrophobic assembly". *Nature* 437 (2005), pp. 640–647.

[8]J. A. McCammon. "Theory of biomolecular recognition". *Curr. Opin. Struct. Biol.* 8.2 (1998), pp. 245–249.

[9]B. J. Berne, J. D. Weeks, and R. Zhou. "Dewetting and hydrophobic interaction in physical and biological systems". *Annu. Rev. Phys. Chem.* 60 (2009), pp. 85–103.

[10]Y. S. Djikaev and E. Ruckenstein. "Recent developments in the theoretical, simulational, and experimental studies of the role of water hydrogen bonding in hydrophobic phenomena". *Adv. Colloid Interface Sci.* 235 (2016), pp. 23–45.

[11]N. T. Southall, K. A. Dill, and A. D. J. Haymet. "A view of the hydrophobic effect". *J. Phys. Chem. B* 106.3 (2002), pp. 521–533.

[12]C. Y. Lee, J. A. McCammon, and P. J. Rossky. "The structure of liquid water at an extended hydrophobic surface". *J. Chem. Phys.* 80.9 (1984), pp. 4448–4455.

[13]A. Nicholls, K. A. Sharp, and B. Honig. "Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons". *Proteins: Struct., Funct., and Bioinf.* 11.4 (1991), pp. 281–296.

[14]K. Lum, D. Chandler, and J. D. Weeks. "Hydrophobicity at small and large length scales". *J. Phys. Chem. B* 103.22 (1999), pp. 4570–4577.

[15]D. M. Huang and D. Chandler. “Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding”. *Proc. Natl. Acad. Sci. (USA)* 97.15 (2000), pp. 8324–8327.

[16]D. M. Huang, P. L. Geissler, and D. Chandler. “Scaling of hydrophobic solvation free energies”. *J. Phys. Chem. B* 105.28 (2001), pp. 6704–6709.

[17]D. M. Huang and D. Chandler. “The hydrophobic effect and the influence of solute-solvent attractions”. *J. Phys. Chem. B* 106.8 (2002), pp. 2047–2053.

[18]S. Rajamani, T. Ghosh, and S. Garde. “Size dependent ion hydration, its asymmetry, and convergence to macroscopic behavior”. *J. Chem. Phys.* 120.9 (2004), pp. 4457–4466.

[19]S. Rajamani, T. M. Truskett, and S. Garde. “Hydrophobic hydration from small to large lengthscales: Understanding and manipulating the crossover”. *Proc. Natl. Acad. Sci. (USA)* 102.27 (2005), pp. 9475–9480.

[20]H. S. Ashbaugh and L. R. Pratt. “Colloquium: Scaled particle theory and the length scales of hydrophobicity”. *Rev. Mod. Phys.* 78.1 (2006), pp. 159–178.

[21]L. R. Pratt. “Molecular theory of hydrophobic effects: ‘She is too mean to have her name repeated.’ ” *Ann. Rev. Phys. Chem.* 53 (2002), pp. 409–436.

[22]F. Sedlmeier, D. Horinek, and R. R Netz. “Entropy and enthalpy convergence of hydrophobic solvation beyond the hard-sphere limit”. *J. Chem. Phys.* 134.5 (2011), 02B604.

[23]F. Sedlmeier and R. R. Netz. “The spontaneous curvature of the water-hydrophobe interface”. *J. Chem. Phys.* 137.13 (2012), p. 135102.

[24]Y. S. Djikaev and E. Ruckenstein. “Temperature effect on small-to-large crossover lengthscale of hydrophobic hydration”. *J. Chem. Phys.* 139.18 (2013), p. 184709.

[25]J. G. Davis, K. P. Gierszal, P. Wang, and D. Ben-Amotz. “Water structural transformation at molecular hydrophobic interfaces”. *Nature* 491.7425 (2012), pp. 582–585.

[26]H. S. Ashbaugh. “Entropy crossover from molecular to macroscopic cavity hydration”. *Chem. Phys. Lett.* 477.1 (2009), pp. 109–111.

[27]R. G. Weiß, M. Heyden, and J. Dzubiella. “Curvature dependence of hydrophobic hydration dynamics”. *Phys. Rev. Lett* 114.18 (2015), p. 187802.

[28]K. Katsov and J. D. Weeks. “On the mean field treatment of attractive interactions in nonuniform simple fluids”. *J. Phys. Chem B* 105.28 (2001), pp. 6738–6744.

[29]D. M. Huang and D. Chandler. “Cavity formation and the drying transition in the Lennard-Jones fluid”. *Phys. Rev. E* 61.2 (2000), pp. 1501–1506.

[30]A. Jain and H. S. Ashbaugh. “Digging a hole: Scaled-particle theory and cavity solvation in organic solvents”. *J. Chem. Phys.* 129.17 (2008), p. 174505.

[31]E. Wu and S. Garde. “Lengthscale-dependent solvation and density fluctuations in n-octane”. *J. Phys. Chem. B* 119.29 (2015), pp. 9287–9294.

[32]E. Wu and S. Garde. “Understanding n-octane behavior near graphene with scaled solvent–solute attractions”. *J. Phys. Chem. B* 120.8 (2016), pp. 2033–2042.

[33]G. Hummer, S. Garde, A. E. García, A. Pohorille, and L. R. Pratt. “An information theory model of hydrophobic interactions”. *Proc. Natl. Acad. Sci.* 93.17 (1996), pp. 8951–8955.

[34]F. H. Stillinger. “Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory”. *J. Sol. Chem.* 2.2/3 (1973), pp. 141–158.

[35]K. Lum and A. Luzar. “Pathway to surface-induced phase transition of a confined fluid”. *Phys. Rev. E* 56.6 (1997), R6283–R6286.

[36]P. G. Bolhuis and D. Chandler. “Transition path sampling of cavitation between molecular scale solvophobic surfaces”. *J. Chem. Phys.* 113.18 (2000), pp. 8154–8160.

[37]K. Leung, A. Luzar, and D. Bratko. “Dynamics of capillary drying in water”. *Phys. Rev. Lett.* 90.6 (2003), p. 065502.

[38]J. L. Parker, P. M. Claesson, and P. Attard. “Bubbles, cavities, and the long-ranged attraction between hydrophobic surfaces.” *J. Phys. Chem.* 98.34 (1994), pp. 8468–8480.

[39]S. Sharma and P. G. Debenedetti. “Evaporation rate of water in hydrophobic confinement”. *Proc. Natl. Acad. Sci. (USA)* 109.12 (2012), pp. 4365–4370.

[40]C. A. Cerdeiriña, P. G. Debenedetti, P. J. Rossky, and N. Giovambattista. “Evaporation length scales of confined water and some common organic liquids”. *J. Phys. Chem. Lett.* 2.9 (2011), pp. 1000–1003.

[41]H. S. Ashbaugh. “Solvent cavitation under solvophobic confinement”. *J. Chem. Phys.* 139.6 (2013), p. 064702.

[42]Y. E. Altabet and P. G. Debenedetti. “The role of material flexibility on the drying transition of water between hydrophobic objects: A thermodynamic analysis”. *J. Chem. Phys.* 141.18 (2014), p. 18C531.

[43]L. Guillemot, T. Biben, A. Galarneau, G. Vigier, and É. Charlaix. “Activated drying in hydrophobic nanopores and the line tension of water”. *Proc. Natl. Acad. Sci. (USA)* 109.48 (2012), pp. 19557–19562.

[44]R. C. Remsing, E. Xi, S. Vembanur, et al. “Pathways to dewetting in hydrophobic confinement”. *Proc. Natl. Acad. Sci. (USA)* 112.27 (2015), pp. 8181–8186.

[45]M. Kanduč and R. R Netz. “From hydration repulsion to dry adhesion between asymmetric hydrophilic and hydrophobic surfaces”. *Proc. Natl. Acad. Sci. (USA)* 112.40 (2015), pp. 12338–12343.

[46]R. Nagarajan and E. Ruckenstein. “Theory of surfactant self-assembly: A predictive molecular thermodynamic approach”. *Langmuir* 7.12 (1991), pp. 2934–2969.

[47]R. Baron, P. Setny, and J. A. McCammon. *Protein-Ligand Interactions*. Weinheim: Wiley-VCH, 2012. Chap. Hydrophobic Association and Volume-Confined Water Molecules.

[48]P. Setny and M. Geller. “Water properties inside nanoscopic hydrophobic pocket studied by computer simulations”. *J. Chem. Phys.* 125.14 (2006), p. 144717.

[49]P. Setny. “Water properties and potential of mean force for hydrophobic interactions of methane and nanoscopic pockets studied by computer simulations”. *J. Chem. Phys.* 127.5 (2007), p. 054505.

[50]P. Setny. “Hydrophobic interactions between methane and a nanoscopic pocket: Three dimensional distribution of potential of mean force revealed by computer simulations”. *J. Chem. Phys.* 128.12 (2008), p. 125105.

[51]L.-T. Cheng, Z. Wang, P. Setny, et al. “Interfaces and hydrophobic interactions in receptor-ligand systems: A level-set variational implicit solvent approach”. *J. Chem. Phys.* 131 (2009), p. 144102.

[52]P. Setny, Z. Wang, L.-T. Cheng, et al. “Dewetting-controlled binding of ligands to hydrophobic pockets”. *Phys. Rev. Lett.* 103 (2009), p. 187801.

[53]P. Setny, R. Baron, and J. A. McCammon. “How can hydrophobic association be enthalpy driven?” *J. Chem. Theory Comput.* 6.9 (2010), pp. 2866–2871.

[54]R. Baron, P. Setny, and J. A. McCammon. “Water in cavity-ligand recognition”. *J. Am. Chem. Soc.* 132.34 (2010), pp. 12091–12097.

[55]P. Setny, R. Baron, P. Kekenes-Huskey, J. A. McCammon, and J. Dzubiella. “Solvent fluctuations in hydrophobic cavity-ligand binding kinetics”. *Proc. Natl. Acad. Sci. (USA)* 110.4 (2013), pp. 1197–1202.

[56]T. Young, R. Abel, B. Kim, B. J. Berne, and R. A. Friesner. “Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding”. *Proc. Natl. Acad. Sci. (USA)* 104.3 (2007), pp. 808–813.

[57]S. K. Nair, T. L. Calderone, D. M. Christianson, and C. A. Fierke. “Altering the mouth of a hydrophobic pocket – Structure and kinetics of human carbonic anhydrase-II mutants at residue VAL-121”. *J. Phys. Chem. B* 266.26 (1991), pp. 17320–17325.

[58]C. Carey, Y.-K. Cheng, and P. J. Rossky. “Hydration structure of the α-chymotrypsin substrate binding pocket: the impact of constrained geometry”. *Chem. Phys.* 258.2 (2000), pp. 415–425.

[59]F. Biedermann, W. M. Nau, and H.-J. Schneider. “The hydrophobic effect revisited – Studies with supramolecular complexes imply high-energy water as a noncovalent driving force”. *Angew. Chem.* 53.42 (2014), pp. 11158–11171.

[60]Q. Shi, M. P. Mower, D. G. Blackmond, and J. Rebek Jr. “Water-soluble cavitand promote hydrolyses of long-chain diesters”. *Proc. Natl. Acad. Sci. (USA)* 113.33 (2016), pp. 9199–9203.

[61]M. B. Hillyer and B. C. Gibb. “Molecular shape and the hydrophobic effect”. *Annu. Rev. Phys. Chem.* 67 (2016), pp. 307–329.

[62]P. P. Wanjari, B. C. Gibb, and H. S. Ashbaugh. “Simulation optimization of spherical non-polar guest recognition by deep-cavity cavitands”. *J. Chem. Phys.* 139.23 (2013), p. 234502.

[63]Y. Hu, D. Devegowda, A. Striolo, et al. “The dynamics of hydraulic fracture water confined in nanopores in shale reservoirs”. *J. Unconv. Oil Gas Resour.* 9 (2015), pp. 31–39.

[64]A. Phan, D. R. Cole, and A. Striolo. “Aqueous methane in slit-shaped silica nanopores: high solubility and traces of hydrates”. *J. Phys. Chem. C* 114.9 (2014), pp. 4860–4868.

[65]A. Phan, A. R. Cole, R. G. Weiß, J. Dzubiella, and A. Striolo. “Confined Water Determines Transport Properties of Guest Molecules in Narrow Pores”. *ACS Nano* 10.8 (2016), pp. 7646–7656.

[66]G. D. Rose, A. R. Geselowitz, G. J. Lesser, R. H. Lee, and M. H. Zehfus. “Hydrophobicity of amino acid residues in globular proteins”. *Science* 229.4716 (1985), pp. 834–828.

[67]K. A. Dill. “Dominant forces in protein folding”. *Biochemistry* 29.31 (1990), pp. 7133–7155.

[68]Walter Kauzmann. “Some factors in the interpretation of protein denaturation”. *Adv. Protein Chem.* 14 (1959), pp. 1–63.

[69]C. Camilloni, D. Bonetti, A. Morrone, et al. “Towards a structural biology of the hydrophobic effect in protein folding”. *Sci. Rep.* 6.28285 (2016), pp. 1–9.

[70]S. H. Gellman. “Introduction: Molecular Recognition”. *Chem. Rev.* 97.5 (1997), pp. 1231–1232.

[71]C. U. Kim, W. Lew, M. A. Williams, et al. “Influenza Neuraminidase Inhibitors Possessing a Novel Hydrophobic Interaction in the Enzyme Active Site. Design, Synthesis, and Structural Analysis of Carbocyclic Sialic Acid Analogues with Potent Anti-Influenza Activity”. *J. Am. Chem. So.* 119.4 (1997), pp. 681–690.

[72]R. C. Bernardi, I. Cann, and K. Schulten. “Molecular dynamics study of enhanced Man5B enzymatic activity”. *Biotechnol. Biofuels* 7.1 (2014), p. 83.

[73]V. B. Sheshanath, J. L. Steven, and V. Bhosale Sidhanath. “Comparative binding study of neurotransmitters in hydrophobic and hydrophilic yoctowells in water”. *Supramol. Chem.* 21.1/2 (2009), pp. 18–23.

[74]S. Borngraeber, M.-J. Budny, G. Chiellini, et al. “Ligand selectivity by seeking hydrophobicity in thyroid hormone receptor”. *Proc. Natl. Acad. Sci. USA* 100.26 (2003), pp. 15358–15363.

[75]C. Hansch, J. P. Björkroth, and A. Leo. “Hydrophobicity and central nervous system agents: On the principle of minimal hydrophobicity in drug design”. *J. Pharm. Sci.* 76.9 (1987), pp. 663–687.

[76]M. B. Hillyer, C. L. D. Gibb, P. Sokkalingam, et al. “Synthesis of water-soluble deep-cavity cavitands”. *Org. Lett.* 18.16 (2016), pp. 4048–4051.

[77]M. Xue, Y. Yang, X. Chi, Z. Zhang, and F. Huang. “Pillararenes, a new class of macrocycles for supramolecular chemistry”. *Acc. Chem. Res.* 45.8 (2012), pp. 1294–1308.

[78]D. E. Koshland. “The key-lock theory and the induced fit theory”. *Angew. Chem., Int. Ed. Engl.* 33.23-24 (1995), pp. 2375–2378.

[79]J. Hu, Y. Cheng, Q. Wu, L. Zhao, and T. Xu. “Host-guest chemistry of dendrimer-drug complexes. 2. Effects of molecular properties of guests and surface functionalities of dendrimers”. *J. Phys. Chem. B* 113.31 (2009), pp. 10650–10659.

[80]S. Sacanna, W. T. M. Irvine, P. M. Chaikin, and D. J. Pine. “Lock and key colloids”. *Nature* 464.7288 (2010), pp. 575–578.

[81]P. Ball. “Water as an active constituent in cell biology”. *Chem. Rev.* 108.1 (2008), pp. 74–108.

[82]J. R. Tame, G. N. Murshudov, E. J. Dodson, et al. “The structural basis of sequence-independent peptide binding by OppA protein”. *Science* 264.5165 (1994), pp. 1578–1581.

[83]J. R. Tame, S. H. Sleigh, A. J. Wilkinson, and J. E. Ladbury. “The role of water in sequence-independent ligand binding by an oligopeptide transport protein”. *Nat. Struct. Biol.* 3.12 (1996), pp. 998–1001.

[84]E. Chung, D. Henriques, D. Renzoni, et al. “Mass spectrometric and thermodynamic studies reveal the role of water molecules in complexes formed between SH2 domains and tyrosyl phosphopeptides”. *Structure* 6.9 (1998), pp. 1141–1151.

[85]F. A. Quiocho, D. K. Wilson, and N. K. Vyas. “Substrate specificity and affinity of a protein modulated by bound water molecules”. *Nature* 340.6232 (1989), pp. 404–407.

[86]P. W. Snyder, M. R. Lockett, D. T. Moustakas, and G. M. Whitesides. “Is it the shape of the cavity, or the shape of the water in the cavity?” *Eur. Phys. J. Special Topics* 223.5 (2014), pp. 853–891.

[87]T. Young, L. Hua, X. Huang, et al. “Dewetting transitions in protein cavities”. *Proteins* 78.8 (2010), pp. 1856–1869.

[88]J. D. Dunitz. “Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions”. *Chem. Biol.* 2.11 (1995), pp. 709–712.

[89]J. Dzubiella. “How interface geometry dictates water’s thermodynamic signature in hydrophobic association”. *J. Stat. Phys.* 145.2 (2011), pp. 227–239.

[90]N. Giovambattista, C. F. Lopez, P. J. Rossky, and P. G. Debenedetti. “Hydrophobicity of protein surfaces: separating geometry from chemistry”. *Proc. Natl. Acad. Sci. (USA)* 105.7 (2007), pp. 2274–2279.

[91]Johan Qvist, Monica Davidovic, Donald Hamelberg, and Bertil Halle. “A dry ligand-binding cavity in a solvated protein”. *Proc. Natl. Acad. Sci. (USA)* 105.17 (2008), pp. 6296–6301.

[92]Subramanian Vaitheeswaran, Hao Yin, Jayendran C. Rasaiah, and Gerhard Hummer. “Water clusters in nonpolar cavities”. *Proc. Natl. Acad. Sci. (USA)* 101.49 (2004), pp. 17002–17005.

[93]Y.-K. Cheng and P. J. Rossky. “Surface topography dependence of biomolecular hydrophobic hydration”. *Nature* 392.6677 (1998), pp. 696–699.

[94]Y.-K. Cheng and P. J. Rossky. “The effect of vicinal polar and charged groups on hydrophobic hydration”. *Biochemistry* 50.7 (1999), pp. 742–750.

[95]K. Haider, L. Wickstrom, S. Ramsey, M. K. Gilson, and T. Kurtzman. “Enthalpic breakdown of water structure on protein active-site surfaces”. *J. Phys. Chem. B* 120.34 (2016), pp. 8743–8756.

[96]A. C. Pan, D. W. Borhani, R. O. Dror, and D. E. Shaw. “Molecular determinants of drug-receptor binding kinetics”. *Drug Discov. Today* 18.13-14 (2013), pp. 667–673.

[97]D. C. Swinney. “The role of binding kinetics in therapeutically useful drug action”. *Curr. Opin. Drug Discov. Dev.* 12.1 (2009), pp. 31–39.

[98]R. A. Copeland, Pompliano D. L., and Meek T. D. “Drug-target residence time and tis implications for lead optimization”. *Nat. Rev. Drug Discov.* 5.9 (2006), pp. 730–739.

[99]D. Guo, T. Mulder-Krieger, A. P. IJzerman, and L. H. Heitman. “Functional efficacy of adenosine A2A receptor agonists is positively correlated to their receptor residence time”. *Br. J. Pharmacol.* 166.6 (2012), pp. 1846–1859.

[100]P. Ball. “Water is an active matrix of life for cell and molecular biology”. *Proc. Natl. Acad. Sci. (USA)* (2017). DOI: 10.1073/pnas.1703781114.

[101]J. Mondal, J. A. Morrone, and B. J. Berne. “How hydrophobic drying forces impact the kinetics of molecular recognition”. *Proc. Natl. Acad. Sci. (USA)* 110.33 (2013), pp. 13277–13282.

[102]J. Li, J. A. Morrone, and B. J. Berne. “Are hydrodynamic interactions important in the kinetics of hydrophobic collapse?” *J. Phys. Chem. B* 116.37 (2012), pp. 11537–11544.

[103]J. A. Morrone, J. Li, and B. J. Berne. “Interplay between hydrodynamics and the free energy surface in the assembly of nanoscale hydrophobes”. *J. Phys. Chem. B* 116.1 (2012), pp. 378–389.

[104]P. Tiwary, J. Mondal, J. A. Morrone, and B. J. Berne. “Role of water and steric constraints in the kinetics of cavity–ligand unbinding”. *Proc. Natl. Acad. Sci. (USA)* 112.39 (2015), pp. 12015–12019.

[105]P. Tiwary and B. J. Berne. “How wet should be the reaction coordinate for ligand unbinding?” *J. Chem. Phys.* 145.5 (2016).

[106]J. Mondal, R. A. Friesner, and B. J. Berne. “Role of desolvation in thermodynamics and kinetics of ligand binding to a kinase”. *J. Chem. Theory Comput.* 10.12 (2014), pp. 5696–5705.

[107]L. R. Pratt and D. Chandler. “Theory of the hydrophobic effect”. *J. Chem. Phys.* 67.8 (1977), pp. 3683–3704.

[108]R. Brown. “XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies”. *Philosophical Magazine Series 2* 4.21 (1828), pp. 161–173.

[109]A. Einstein. “Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen”. *Ann. Phys. (Berlin, Ger.)* 322 (1905), pp. 549–560.

[110]P. Langevin. “Sur la théorie du mouvement brownien”. *CR Acad. Sci. Paris* 146 (1908), pp. 520–533.

[111]N. Pottier. *Nonequilibirum statistical physics: linear irreversible processes*. Oxford University Press, 2010.

[112]D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. Academic press, 2001.

[113]M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, 1987.

[114]B. Hess, H. Bekker, and J. G. E. M. Fraaije H. J. C. Berendsen. “LINCS: A linear constraint solver for molecular simulations”. *J. Comput. Chem.* 18.12 (1997), pp. 1463–1472.

[115]J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. “Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes”. *J. Comp. Phys.* 23.3 (1977), pp. 327–341.

[116]S. Miyamoto and P. A. Kollman. “SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models”. *J. Comput. Chem.* 13.8 (1992), pp. 952–962.

[117]S. Pronk, S. Páll, R. Schulz, et al. “GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit”. *Bioinformatics* 29.7 (2013), pp. 845–854.

[118]M. J. Abraham, T. Murtola, R. Schulz, et al. “GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers”. *SoftwareX* 1 (2015), pp. 19–25.

[119]J. D. Bernal and R. H. Fowler. “A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions”. *J. Chem. Phys.* 1.8 (1933), pp. 515–548.

[120]J. A. Barker and R. O. Watts. “Structure of water; A Monte Carlo calculation”. *Chem. Phys. Lett.* 3.3 (1969), pp. 144–145.

[121]A. Rahman and F. H. Stillinger. “Molecular dynamics study of liquid water”. *J. Chem. Phys.* 55.7 (1971), pp. 3336–3359.

[122]J. F. Ouyang and R. P. Bettens. “Modelling water: a lifetime enigma”. *Chimia* 69.3 (2015), pp. 104–111.

[123]C. Millot and A. J. Stone. “Towards and accurate intermolecular potential for water”. *Mol. Phys.* 77.3 (1992), pp. 439–462.

[124]X. Huang, B. J. Braams, and J. M. Bowman. “Ab initio potential energy and dipole moment surfaces of $(H_2O)_2$”. *J. Phys. Chem. A* 110.2 (2006), pp. 445–451.

[125]R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird. “Interaction potential for water dimer from symmetry-adapted perturbation theory based on density functional description of monomers”. *J. Chem. Phys.* 125.4 (2006).

[126]H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. “The missing term in effective pair potentials”. *J. Phys. Chem.* 91.24 (1987), pp. 6269–6271.

[127]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. “Comparison of simple potential functions for simulating liquid water”. *J. Chem. Phys.* 79.2 (1983), pp. 926–935.

[128]W. L. Jorgensen, J. D. Madura, and C. J. Swenson. “Optimized intermolecular potential functions for liquid hydrocarbons”. *J. Am. Chem. Soc.* 106.22 (1984), pp. 6638–6646.

[129]W. L. Jorgensen, D. Maxwell, and J. Tirado-Rives. “Develoment and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids”. *J. Am. Chem. Soc.* 118.45 (1996), pp. 11225–11236.

[130]S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method". *J. Comput. Chem.* 13.8 (1992), pp. 1011–1021.

[131]F. Zhu and G. Hummer. "Convergence and error estimation in free energy calculations using the weighted histogram analysis method". *J. Comput. Chem.* 33.4 (2012), pp. 453–465.

[132]G. Hummer. "Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations". *New J. Phys.* 7.1 (2005), p. 34.

[133]J. Mittal and G. Hummer. "Static and dynamic correlations in water at hydrophobic interfaces". *Proc. Natl. Acad. Sci. (USA)* 105.51 (2008), pp. 20130–20135.

[134]A. J. Patel, P. Varilly, S. N. Jamadagni, et al. "Sitting at the edge: How biomolecules use hydrophobicity to tune their interactions and function". *J. Phys. Chem. B* 116.8 (2012), pp. 2498–2503.

[135]A. J. Patel, P. Varilly, and D. Chandler. "Fluctuations of water near extended hydrophobic and hydrophilic surfaces". *J. Phys. Chem. B* 114.4 (2010), pp. 1632–1637.

[136]S. N. Jamadagni, R. Godawat, and S. Garde. "Hydrophobicity of proteins and interfaces: Insights from density fluctuations". *Annu. Rev. Chem. Biomol. Eng.* 2 (2011), pp. 147–171.

[137]E. Gianti, L. Delemotte, M. L. Klein, and V. Carnevale. "On the role of water density fluctuations in the inhibition of a proton channel". *Proc. Natl. Acad. Sci. (USA)* 113.52 (2016), E8359–E8368.

[138]R. Remsing and A. J. Patel. "Water density fluctuations relevant to hydrophobic hydration are unaltered by attractions". *J. Chem. Phys.* 142.2 (2015), p. 024502.

[139]E. Xi, R. C. Remsing, and A. J. Patel. "Sparse sampling of water density fluctuations in interfacial environments". *J. Chem. Theory Comput.* 12.2 (2016), pp. 706–713.

[140]R. Underwood and D. Ben-Amotz. "Communication: Length scale dependent oil-water energy fluctuations". *J. Chem. Phys.* 135.20 (2011), p. 201102.

[141]P. Hänggi and P. Talkner. "Non-Markov processes: The problem of the mean first passage time". *Z. Phys. B - Condensed Matter* 45.1 (1981), pp. 79–83.

[142]P. Hänggi and P. Talkner. "First-passage time problems for non-Markovian processes". *Phys. Rev. A* 32.3 (1985), pp. 1934–1937.

[143]P. V. E. Moss F. an McClintock. *Theory of continuous Fokker-Planck systems*. Cambridge University Press, 1989, pp. 112–114.

[144]A. J. F. Siegert. “On the first passage time probability problem”. *Phys. Rev.* 81.4 (1951), pp. 617–623.

[145]G. H. Weiss. “First passage time problems in chemical physics”. *Adv. Chem. Phys.* 13.1 (2007), pp. 1–18.

[146]M. Hinczewski, Y. von Hansen, J. Dzubiella, and R. R. Netz. “How the diffusivity profile reduces the arbitrariness of protein folding free energies”. *J. Chem. Phys.* 132.24 (2010), p. 245103.

[147]F. Seldmeier, Y. von Hansen, L. Mengyu, D. Horinek, and R. R. Netz. “Water dynamics at interfaces and solutes: Disentangling free energy and diffusivity contributions”. *J. Stat. Phys.* 145.2 (2011), pp. 240–252.

[148]P. E. Smith and W. F. van Gunsteren. “The viscosity of SPC and SPC/E water at 277 and 300 K”. *Chem. Phys. Lett.* 215.4 (1993), pp. 315–318.

[149]C. Vega and E. de Miquel. “Surface tension of the most popular models of water by using the test-area simulation method”. *J. Chem. Phys.* 126.15 (2007), p. 154707.

[150]P. Hänggi, P. Talkner, and M. Borkovec. “Reaction-rate theory: Fifty years after Kramers”. *Rev. Mod. Phys.* 62.2 (1990), pp. 251–341.

[151]D. L. Ermak and J. A. McCammon. “Brownian dynamics with hydrodynamic interactions”. *J. Chem. Phys.* 69.4 (1978), pp. 1352–1360.

[152]A. M. Ferrenberg and R. H. Swendsen. “Optimized Monte Carlo data analysis”. *Phys. Rev. Lett.* 63.12 (1989), pp. 1195–1198.

[153]P. Hänggi and P. Talkner. “Memory index of the first passage time: A simple measure of non-Markovian character”. *Phys. Rev. Lett.* 51.25 (1983), pp. 2242–2245.

[154]H. Risken. *The Fokker-Planck equation: Methods of solution and application*. Springer, Berlin, 1996, pp. 179–195.

[155]N. G. van Kampen. *Stochastic processes in physics and chemistry*. North Holland, Amsterdam, 2007, pp. 240–242.

[156]E. Pollak and A. M. Berezhkovsii. “Fokker-Planck equation for nonlinear stochastic dynamics in the presence of space and time dependent friction”. *J. Chem. Phys.* 99.2 (1993), pp. 1344–1346.

[157]A.M. Berezhkovsii, A.M. Frishman, and E. Pollak. “Variational transition state theory for multidimensional activated rate processes in the presence of anisotropic friction”. *J. Chem. Phys.* 101.6 (1994), pp. 4778–4789.

[158]P. J. Cregg, D. S. F. Crothers, and A. W. Wickstead. “An approximate formula for the relaxation time of a single domain ferromagnetic particle with uniaxial anisotropy and collinear field”. *J. Appl. Phys.* 76.8 (1994), pp. 4900–4902.

[159]Yu. P. Kalmykov, W. T. Coffey, and J. T. Waldron. “Exact analytic solution for the correlation time of a Brownian particle in a double-well potential from the Langevin equation”. *J. Chem. Phys.* 105.5 (1996), pp. 2112–2118.

[160]R. G. Weiß, P. Setny, and J. Dzubiella. “Solvent fluctuations induce non-Markovian kinetics in hydrophobic pocket-ligand binding”. *J. Phys. Chem. B* 120.33 (2016), pp. 8127–8136.

[161]H. C. Hamaker. “The London – van der Waals attraction between spherical particles”. *Physica* 4.10 (1937), pp. 1058–1072.

[162]M.-C. Bellissent-Funel, A. Hassanali, M. Havenith, et al. “Water determines the structure and dynamics of proteins”. *Chem. Rev.* 116.13 (2016), pp. 7673–7697.

[163]K. Meister, S. Ebbinghaus, Y. Xu, et al. “Long-range protein–water dynamics in hyperactive insect antifreeze proteins”. *Proc. Natl. Acad. Sci. (USA)* 110.5 (2013), pp. 1617–1622.

[164]S. Ebbinghaus, S. J. Kim, M. Heyden, et al. “An extended dynamical hydration shell around proteins”. *Proc. Natl. Acad. Sci. (USA)* 104.52 (2007), pp. 20749–20752.

[165]J. G. Kirkwood and J. B. Shumaker. “Forces between protein molecules in solution arising from fluctuations in proton charge and configuration”. *Proc. Natl. Acad. Sci. (USA)* 38.10 (1952), pp. 863–871.

[166]J. G. Kirkwood and J. B. Shumaker. “The influence of dipole moment fluctuations on the dielectric increment of proteins in solution”. *Proc. Natl. Acad. Sci. (USA)* 38.10 (1952), pp. 855–862.

[167]Alexander M Berezhkovskii and Sergey M Bezrukov. “Effect of stochastic gating on channel-facilitated transport of non-interacting and strongly repelling solutes”. *The Journal of Chemical Physics* 147.8 (2017), p. 084109.

[168]A. Szabo, D. Shoup, S. H. Northrup, and J. A. McCammon. “Stochastically gated diffusion-influenced reactions”. *J. Chem. Phys.* 77.9 (1982), pp. 4484–4493.

[169]N. Agmon and J. J. Hopfield. “Transient kinetics of chemical reactions with bounded diffusion perpendicular to the reaction coordinate: intramolecular processes with slow conformational changes”. *J. Chem. Phys.* 78.11 (1983), pp. 6947–6959.

[170]P. R. Batista., G. Pandey, P. G. Pascutti, et al. “Free energy profiles along consensus normal modes provide insight into HIV-1 protease flap opening”. *J. Chem. Theory Comput.* 7.8 (2011), pp. 2348–2352.

[171]V. Srajer, L. Reinisch, and P. M. Champion. “Protein fluctuations, distributed coupling, and the binding of ligands to heme proteins”. *J. Am. Chem. Soc.* 110.20 (1988), pp. 6656–6670.

[172]H.-X. Zhou. “From induced fit to conformational selection: A continuum of binding mechanism controlled by the timescale of conformational transitions”. *Biophys. J.* 98.6 (2010), pp. L15–L17.

[173]L. Cai and H.-X. Zhou. “Theory and simulation on the kinetics of protein-ligand binding coupled to conformational change”. *J. Chem. Phys.* 134.10 (2011), p. 105101.

[174]D. Beece, L. Eisenstein, H. Frauenfelder, et al. “Solvent viscosity and protein dynamics”. *Biochemistry* 19.23 (1980), pp. 5147–5157.

[175]W. Doster. “Viscosity scaling and protein dynamics”. *Biophys. Chem.* 17.2 (1983), pp. 97–103.

[176]H. Frauenfelder and P. G. Wolynes. “Rate theories and puzzles of hemeprotein kinetics”. *Science* 229 (1985), pp. 337–345.

[177]P. Hänggi. “Physics of ligand migration in biomolecules”. *J. Stat. Phys.* 30.2 (1983), pp. 401–412.

[178]N. Pizzolato, A. Fiasconaro, D. P. Adorno, and B. Spagnolo. “Resonant activation in polymer translocation: new insights into the escape dynamics of molecules driven by an oscillating field”. *Phys. Biol.* 7.3 (2010), p. 034001.

[179]W. K. Kim, C. Hyeon, and W. Sung. “Weak temporal signals can synchronize and accelerate the transition dynamics of biopolymers under tension”. *Proc. Natl. Acad. Sci. (USA)* 109.36 (2012), pp. 14410–14415.

[180]S. Wu, J. Dzubiella, J. Kaiser, et al. “Thermosensitive Au-PNIPA Yolk–Shell Nanoparticles with Tunable Selectivity for Catalysis”. *Angew. Chem., Int. Ed.* 51.9 (2012), pp. 2229–2233.

[181]S. Angioletti-Uberti, Y. Lu, M. Ballauff, and J. Dzubiella. “Theory of solvation-controlled reactions in stimuli-responsive nanoreactors”. *J. Phys. Chem. C* 119.27 (2015), pp. 15723–15730.

[182]J. J. Kolb, S. Angioletti-Uberti, and J. Dzubiella. “Communication: Resonance reaction in diffusion-influenced bimolecular reactions”. *J. Chem. Phys.* 144.8 (2016), p. 081102.

[183]R. Roa, W. K. Kim, M. Kanduč, J. Dzubiella, and S. Angioletti-Uberti. “Catalyzed bimolecular reactions in responsive nanoreactors”. *ACS Catal.* 7.9 (2017), pp. 5604–5611.

[184]C. R. Doering and J. C. Gadoua. “Resonant activation over a fluctuating barrier”. *Phys. Rev. Lett.* 69.16 (1992), p. 2318.

[185]P. Liu, X. Huang, R. Zhou, and B. J. Berne. “Observation of dewetting transition in the collapse of the melittin tetramer”. *Nature* 437.7055 (2005), pp. 159–162.

[186]J. Masoliver, K. Lindenberg, and B. J. West. “First-passage times for non-Markovian processes”. *Phys. Rev. A* 33.3 (1986), pp. 2177–2180.

[187]J. M. Sancho, F. Sagués, and M. San Miguel. “Mean first-passage time of continuous non-Markovian processes driven by colored noise”. *Phys. Rev. A* 33.5 (1986), pp. 3399–3403.

[188]R. Zwanzig. “Ensemble method in the theory of irreversibility”. *J. Chem. Phys.* 33.5 (1960), pp. 1338–1341.

[189]R. H. Terwiel. “Projection operator method applied to stochastic linear differential equations”. *Physica* 74.2 (1974), pp. 248–265.

[190]J. S. Langer. “Statistical theory of decay of metastable states”. *Ann. Phys.* 54.2 (1969), pp. 258–275.

[191]A. M. Berezhkovsii and A. Szabo. “Time scale separation leads to position-dependent diffusion along a slow coordinate”. *J. Chem. Phys.* 135.7 (2011), p. 074108.

[192]G. Hummer and A. Szabo. “Optimal dimensionality reduction of multistate kinetic and Markov-State Models”. *J. Phys. Chem. B* 119.29 (2015), pp. 9029–9037.

[193]P. Tiwary, V. Limongello, M. Salvalaglio, and M. Parrinello. “Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps”. *Proc. Natl. Acad. Sci. (USA)* 112.5 (2014), E386–E391.

[194]G. R. Bowman, V. S. Pande, and F. Noe. *An introduction to markov state models and their application to long timescale molecular simulation*. Dortrecht, Netherlands: Springer, 2014.

[195]A. J. Patel and Garde S. “Efficient method to characterize the context-dependent hydrophobicity of proteins”. *J. Phys. Chem. B* 118.6 (2014), pp. 1564–1573.

[196]F. Sedlmeier, D. Horinek, and R. R. Netz. “Nanoroughness, intrinsic density profile, and rigidity of the air-water interface”. *Phys. Rev. Lett.* 103.13 (2009), p. 136102.

[197]K. Falk and K. Mecke. “Capillary waves of compressible fluids”. *J. Phys.: Condens. Matter* 23.18 (2011), p. 184103.

[198]K. Rane and N. F. A. van der Vegt. “Understanding the influence of capillary waves on solvation at the liquid-vapor interface”. *J. Chem. Phys.* 144.11 (2016), p. 114111.

[199]S. Zhou, L.-T. Cheng, H. Sun, et al. “LS-VISM: A software package for analysis of biomolecular solvation”. *J. Comput. Chem.* 36.14 (2015), pp. 1047–1059.

[200]S. Zhou, H. Sun, L.-T. Cheng, et al. “Stochastic level-set variational implicit-solvent approach to solute-solvent interfacial fluctuations”. *J. Chem. Phys.* 145.5 (2016), p. 054114.

[201] S. Vaikuntanathan, G. Rotskoff, A. Hudson, and P. L. Geissler. “Necessity of capillary modes in a minimal model of nanoscale hydrophobic solvation”. *Proc. Natl. Acad. Sci. (USA)* 113.16 (2016), E2224–E2230.

Acknowledgement

I want to express my deepest gratitude to Prof. Dr. Joachim Dzubiella who has been my mentor for more than five years. I am very grateful for his support during my undergraduate studies and for giving me the chance to continue my research as a Ph.D. student. His enthusiasm for excellent science praxis pushed me to give my best, while he still gave me the freedom to develop my scientific creativity. Thanks to him I was able to enjoy the scientific discourse on several specialized meetings in places all over the world such as Telluride CO, Holderness NH, and Erice on Sicily, just to mention a few. Last but not least, it has been a great pleasure being supervised in his considerate and inspiring manner with high professional and personal esteem.

I also owe special thank to Dr. Piotr Setny who has been a constant scientific contact. If it was not for his MD model, I would not have been doing this work. His thorough help has been precious all along for our neat collaborative efforts. Meeting him in Berlin, Dresden, or Warsaw has always been a great pleasure.

My thanks also go to all of my colleagues and friends from the Soft Matter Institute at HZB for creating a great scientific atmosphere that is filled with enthusiasm and fun. I will always look back with joy to the time I shared with you.

Special thanks go to Richard Chudoba, Dr. Karol Palczynski, Dr. Matej Kanduč, and Dr. Rafael Roa for careful proofreading of this manuscript and their constructive comments. Moreover, I very much thank Philip Bolton Jr. for applying his professional expertise when he stylistically reviewed this manuscript. Besides, I am proud to call him a friend who ever recurs in my life.

Finally, I want to thank my family and good friends for their patience and understanding throughout this time. I thank my deceased grandparent Dr. Hannelore and Reinhold Petermann who have been my role models. But most of all I thank Honorata – my partner in life and the mother of my son. She eases my soul to calm me from stress and kicks my butt when I need to stand up.

Schlusserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die hier angegebenen Quellen und Hilfsmitteln benutzt habe.

Ferner erkläre ich, dass ich nicht anderweitg mit oder ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich einer Doktorprüfung zu unterziehen.

Berlin, 14. September 2017

Richard Gregor Weiß